FINAL REPORT

# MAPPING RICE PRODUCTION IN CHINA WITH AVHRR IMAGERY

PRINCIPAL INVESTIGATOR: WILLIAM EMERY

NAG 5-3847

This thesis entitled:
Coupling AVHRR Imagery with Biogeochemical Models of Methane Emission from Rice Crops
written by Eleni. J. Paliouras
has been approved for the Department of Aerospace Engineering Science

_____

Dr. William Emery

_____

Dr. James Maslanik

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Paliouras, Eleni. J. (Ph.D., Aerospace Engineering Science)

Coupling AVHRR Imagery with Biogeochemical Models of Methane Emission from Rice Crops

Thesis directed by Dr. William Emery

Rice is a staple food source for much of the world and most of it is grown in paddies which remain flooded for a large part of the growing season. This anaerobic environment is ideal for the activities of methanogenic bacteria, that are responsible for the production of methane gas, some of which is released into the atmosphere. In order to better understand the role that rice cropping plays in the levels of atmospheric methane, several models have been developed to predict the methane flux from the paddies. These models generally utilize some type of nominal plant growth curve based on one or two pieces of ground truth data. Ideally, satellite data could be used instead to provide these models with an estimate of biomass change over the growing season, eliminating the need for related ground truth. A technique proposed to accomplish this is presented here, and results that demonstrate its success when applied to rice cropping areas of Texas are discussed. Also presented is a method for utilizing satellite data to map rice cropping areas that could eventually aid in a scheme for populating a GIS-type database with information on exact rice cropping areas. Such a database could then be directly tied to the methane emission models to obtain flux estimates for extensive regional areas.

## Dedication

To my family, whose love and support has always sustained me.

## Acknowledgements

I would first like to thank Dr. William Emery and to express my deepest appreciation for the financial support and his moral support of me through this process of tremendous learning. Without his encouragement, this project would have likely had a much different outcome. Thanks also go to the rest of my committee members who have always been willing to meet with me to talk about my research and life in general. Sincere thanks go also to the folks at NASA-Goddard who were responsible for funding this project.

The list of people who provided help, data, advice, and guidance to me through the years of working on this project is extensive, and I hope that I do not neglect to mention anyone. I first thank Dr. Yao Huang and Dr. Peter van Bodegom for providing their models to me and for patiently answering all of my questions, big and small. Thanks to Dr. Guowei Wu and Dr. Ted Wilson of TAMU for providing the rice biomass data, and to Dr. Ron Sass of Rice University for the methane flux data. Thanks also to Dr. Anna McClung, Jaynen Cockrell, and Dr. James Stansel of the Texas A&M Beaumont Research Station for their help in collecting data and for their willingness to help me with other information needed in my research. Thanks to Mike Shoppa, Randy Epps, and Larry Harbers of the Lakeside Water Division of LCRA for their willingness to provide maps of rice cropping areas and for allowing me to ride along with a water boss to see the crops first-hand. Thanks to Dr. Sam Feagley and Dr. Tom Hallmark of TAMU for

soils information, and to Dr. Melba Crawford and Solar Smith of the University of Texas for AVHRR data.

Deepest appreciation to Dr. Chuck Fowler for everything he has done to help me over the years and to Dan Baldwin for writing and maintaining the code used to process all of the AVHRR data, and for answering zillions of questions about it all. Thanks to Scott Bacoms for being a good and faithful help to me in writing scripts to do everything. Thanks to my fellow classmates for helping me through the non-thesis portion of my degree (and life in general) through their humor, and their willingness to talk and listen.

Finally, although I will never be able express the real level of my appreciation for the love and support he has shown for me, my warmest thanks to Christoph... danke schatz.

# Contents

**Chapter**

# Tables

**Table**

# Figures

**Figure**

# Chapter 1

## Introduction and Background

The study of rice agriculture is necessary for both the importance of rice as a vital food source and because of the fact that cultivating it has an unfortunate byproduct, namely methane gas. As a food source, rice is a staple for a large majority of the world's population, especially in Asia. Because the populations of many Asian nations are increasing at rapid rates, the production of rice will need to similarly increase. In 1989, it was estimated that the demand for rice would increase by 65% by the year 2019 [26]. Rice crops are considered to be one of the primary anthropogenic sources of methane gas [3],[5], [18], [25], and [43]. A reason for concern is that this gas is a so-called "greenhouse" trace gas and given its increasing levels in the atmosphere, is thought to contribute to the suspected global warming phenomenon. Some estimate that methane may contribute up to 20% to the global warming effect [34].

Trace gas emissions from anthropogenic sources is an issue that generates great worldwide interest because of the fact that mankind is very likely affecting the current and future climate in potentially negative ways. In an effort to better understand these effects, scientists and engineers are conducting research on all of the varied fronts which relate to climate change and biosphere/atmosphere interactions. The study of global warming through increasing concentrations of greenhouse gases is one area which has received much media and scientific attention.

Research fueled by debates on this topic is being conducted on numerous, interrelated fronts in an effort to better understand the complex relationship between human activities and the earth's climate. The research ranges from attempting to verify if the observed data even supports the existence of an anthropogenically generated global-warming phenomenon, to identification of sources and sinks of the trace gases, to measuring the source strengths, to studies which focus on modeling the processes which generate the gases, and finally, to trying to project their impact on the global climatic system [10].

Some of the more commonly known sources of greenhouse gases are related to industry and transportation. Carbon dioxide, $CO_2$, from automobile emissions is one such example. Lesser known are sources from natural and cultivated vegetation, such as the methane, $CH_4$, resulting from rice cropping. While the concentration of atmospheric methane is significantly less than that of carbon dioxide, $CH_4$ has been estimated to have up to 32 times the insulating capability of carbon dioxide, making it an important gas to monitor [26].

The remainder of this chapter will provide some additional background information on the effects of atmospheric methane, and the role that rice agriculture plays as a source in the methane budget. This will be followed by a brief description of efforts to model this source of atmospheric methane. Finally, this chapter will end with a statement of the hypotheses of this thesis, at which time a description of the information contained in the rest of this document will be provided.

## 1.1 Atmospheric Methane

Of the constituent gases in the atmosphere, methane is considered a trace gas because its abundance is very small compared to the major components, as can be noted in Table 1.1.

Table 1.1: Chemical Composition of the Atmosphere - Selected Constituents, adapted from [5]

| Constituent | Chemical Formula | Volume Mixing Ratio (dry air) |
|---|---|---|
| Nitrogen | $N_2$ | 78.084 % |
| Oxygen | $O_2$ | 20.948 % |
| Argon | Ar | 0.934 % |
| Carbon Dioxide | $CO_2$ | 360 ppmv |
| Neon | Ne | 18.18 ppmv |
| Helium | He | 5.24 ppmv |
| **Methane** | $\mathbf{CH_4}$ | **1.7 ppmv** |
| Hydrogen | $H_2$ | 0.55 ppmv |
| Nitrous Oxide | $N_2O$ | 0.31 ppmv |
| Carbon Monoxide | CO | 50-200 ppbv |
| Ozone (tropospheric) | $O_3$ | 10-500 ppbv |
| Ozone (stratospheric) | $O_3$ | 0.5-10 ppm |

Of course, a relatively small amount of a trace gas in the atmosphere does not imply that it can not exert a significant impact on atmospheric chemistry or on the behavior of the atmosphere in general. One commonly talked about trace gas is ozone. Most people are familiar with the hole in the stratospheric ozone layer over Antarctica and the banning of certain chemicals, such as chlorofluorocarbons (CFCs), thought to have caused it. Many people in larger cities are also aware of ozone in the troposphere because there are days when they are warned to minimize outdoor activity because of its harmful effects.

Like ozone, methane is found in both the troposphere and stratosphere, and affects the atmosphere in each through different processes. The effects of an atmospheric constituent can be either direct or indirect. In the case of methane, it is both. The direct effect of atmospheric methane is radiative forcing, which derives from the fact that the spectrum of methane has absorption bands which, in essence, cause radiative energy to become trapped in the atmosphere rather

than being released into space. The indirect effects of methane derive from the oxidation of methane in the atmosphere by hydroxyl, OH, which is the primary sink for methane. In the troposphere, the oxidation of $CH_4$ leads to the formation of $CH_2O$ (formaldehyde), CO, and in the presence of sufficient $NO_x$, ozone. In the stratosphere, oxidation of methane results in $H_2O$ [50]. Thus, the greater the levels of methane in the atmosphere, the greater the decrease in the oxidizing capacity of the atmosphere, and the greater the increase in other undesirable trace gases.

## 1.2 Rice Paddies as a Source of Methane

The largest vegetative source of methane is attributable to the cultivation of rice. Thus, rice paddies and their role as a source of $CH_4$ are important subjects to study. Before discussing the role of rice specifically as a source of atmospheric methane, a brief description of the chemical processes attributable to methane flux from wetland areas in general should come first. These processes can be generalized into: production, consumption, and transport.

The production of methane occurs through a series of microbiological processes which are controlled primarily by the absence of oxygen, such as in flooded waters, and the availability of readily degradable carbon. These two conditions are required for the functioning of the bacteria which produce methane. These methanogenic bacteria, or methanogens, can work only after organic matter has been broken down into usable substrates, such as hydrogen and acetate, by other types of bacteria [35]. The source of the organic matter from which substrates are produced can come from any of the following: mineralizable carbon that is naturally occurring in the soil, the decay of carbon amendments added to the soil, and from the roots of the plants themselves, either by exudation or by decay.

Once methane is produced in the anaerobic soil, some of it is consumed through oxidation. This process is due to activities of methanotrophic bacteria, which require oxygen and, of course, methane. There are two places where there is an oxic/anoxic interface in a flooded environment: at the soil/water interface, and in the rhizosphere, which is the portion of the soil directly under the influence of the root system of a plant [41].

Since water has a low methane solubility, some of the methane produced which is not oxidized by the methanotrophs may eventually be transported from the anaerobic soil layer and released into the atmosphere. There are three transport possibilities: plant-mediated, ebullition of gas bubbles, and diffusion through the soil/water and the water/air interfaces [34]. The debate on the amount of methane that is transported through each of these pathways is far from settled. Additionally, the timing of the methane release(s) is under debate.

Most scientists consider that the majority of the methane flux from rice is derived from plant-mediated transport. In fact, several sources estimate that up to 90% of the methane flux from rice paddies has been transported through the aerenchymal system [14]. The aerenchymal system is an intercellular gas-space system developed in plants that grow in soils that are water-saturated or otherwise deficient in rhizospheric oxygen. In waterlogged soils, there is measurable movement of oxygen from the roots to the rhizosphere, and this same system allows a means for the methane to be transported to the atmosphere [31].

Of course, the goal of this type of research is not to lead to any suggestion that rice cultivation be curtailed. Instead, the goal is to improve estimates of the atmospheric methane budget which would then lead to a better understanding of the global climate. It is through improved understanding of the climate system as a whole, that scientists and governing bodies can more effectively target decreases

for anthropogenic sources of each of the greenhouse gases.

This is not to say that research is not being conducted into ways that rice cropping practices could be changed in order to mitigate the contribution of rice to atmospheric methane levels. On the contrary, many studies have been already been conducted and research is still continuing. It should be recognized that this is not a simple task. Because much of the rice production takes place in low-income countries, in order for $CH_4$ flux mitigation strategies to succeed there can be no loss to the farmer in terms of yield, and no added cost to the farmer in terms of more expensive cropping practices [23]. But there has been some success, especially with research conducted to determine the impact of changing rice cultivar and/or cropping practices on $CH_4$ fluxes. Some of those strategies include simpler ones such as more effective water management and careful selection of soil amendments, and more complex ones such as the breeding of new cultivars which do not exhibit as much potential for methanogenesis [48].

## 1.3 Modeling Methane Emission from Rice Paddies

Because of uncertainties such as those mentioned above on the timing and mechanisms of methane transport, the actual magnitude of rice as a source of atmospheric methane is not precisely known. In an effort to remedy this situation, researchers have begun to model each of the three processes mentioned above (production, consumption, and transport), sometimes as a whole system, sometimes one process at a time. At the start of this research project, four whole-system models were identified as candidates for evaluation. These models were developed by Dr. Mingkui Cao [6], Dr. Yao Huang [17], [15], Dr. Changcheng Li [unpublished], and Dr. Peter van Bodegom [45], [44]. After careful consideration, two of the models were eliminated from consideration for study: the Cao model,

because Dr. Cao is no longer working in the same research area, and the Li model because it is currently undergoing its validation process.

The remaining two models were selected for inclusion in this study. The van Bodegom model, more mechanistic in nature, was developed at Wageningen Agricultural University in the Netherlands. The Huang model, based more on empirical data, was developed at Rice University in Houston, Texas. Although these models will be described in more detail in later sections, it is important to note here that although both of these models consider each of the three processes leading to methane emissions differently, they do have one important thing in common: the need to describe the growth of the rice plant in order to determine the rates of methane production, oxidation, and transport. The plant characteristic that is used to define plant growth in both of these models is total aboveground biomass.

## 1.4 Research Goals and Objectives

As will be described in Chapter 2, data derived from remote sensing satellites has been shown to provide valuable information about vegetation. Because of this, it was recognized that the biomass information required to model methane flux from rice paddies might be derived through the use of satellite remote sensing. In order to test this idea, image data from one satellite sensor, the Advanced Very High Resolution Radiometer (AVHRR), was utilized. The motivation for selecting this particular sensor will be described in Chapter 2.

The geographic focus of the work was the Gulf coast region of Texas, one of the most productive areas of rice cultivation in the world [42]. For this study, AVHRR imagery was acquired for the 1992-1995 and 1998-1999 growing seasons. Plant data from two different test sites in Texas was obtained for these six sea-

sons and the success of the methodology used to relate the AVHRR data and the biomass of rice plants is presented in Chapter 3. The use of biomass values estimated from AVHRR as inputs to the two models will be described in Chapters 4 and 5, and a comparison of the resulting methane emission estimates to methane flux measurements collected in Texas in 1994 will be presented. Chapter 6 will describe a method of detecting rice cropping areas of Texas using AVHRR imagery, and Chapter 7 will provide a summary and present conclusions of the thesis research.

# Chapter 2

# Remote Sensing and the AVHRR Sensor

## 2.1 Introduction

Satellites have been utilized for over forty years to study the earth, its natural phenomena, and mankind's activities upon it. The most obvious benefit of utilizing satellite imagery is that more land area can be studied at one time than can be studied by physically visiting the area of interest. Additionally, the opportunity for repeat coverage of the areas of interest, over long periods of time, is afforded through the use of satellite imagery.

The selection of the satellite sensor from which one wants to use image data depends upon the application of interest. Satellite sensors can be categorized into passive and active sensors. Active imaging sensors utilize their own energy source to illuminate features of interest, whereas passive systems sense naturally occurring emitted or reflected energy [21]. Passive sensors, which make up the majority of systems used today, generally sense wavelengths in the visible and infrared (IR) portions of the electromagnetic spectrum, although there are some passive microwave sensors as well. In general, visible and near infrared (NIR) sensors detect reflected solar radiation, while longer wavelength IR sensors detect emitted thermal radiation. The study of vegetation using remote sensing, which will be discussed in the following section, is most often conducted with sensors

operating in the visible and NIR portions of the spectrum.

## 2.2 Remote Sensing of Vegetation

The study of natural and cropped vegetation is one of the major applications for remote sensing, because satellite data can provide information about how vegetation is changing, if at all, over time and space. The potential for gaining this type of insight is due to knowledge of how reflected solar radiation is altered by the vegetation. As mentioned above, sensors which acquire data in the visible and NIR portions of the spectrum are often utilized, for reasons which will now be described.

Visible channels owe their utility in studying vegetation to the fact that pigments, found primarily in the leaves, absorb energy in this wavelength region. These pigments are responsible for absorbing energy in order to initiate photosynthesis, and are primarily chlorophyll, but also carotene, and xanthophyll. The absorption of visible energy by vegetation is a function of wavelength. In fact, because chlorophyll-a absorbs more in the blue ( $0.4\mu m$-$0.5\mu m$ ) and in the red ( $0.6\mu m$-$0.7\ \mu m$ ) portions of the spectrum, most vegetation appears green [4].

Sensor channels in the NIR region, from $0.75\ \mu m$ to $1.35\ \mu m$, aid in the study of vegetation because the internal structure of leaves causes high levels of reflectance to occur. Basically, because there are no pigments which absorb solar radiation in this wavelength region, the energy enters the internal structure of the leaves where some of it is reflected, and the rest is transmitted through the leaf [24]. Through spectroscopic analysis on individuals leaves, it is known that approximately 50% of the incident radiation is reflected. The reason for this is due to the changes in indices of refraction between the cell walls in the leaves and the air spaces between these cells [4]. These visible and NIR features of the

interaction between solar radiation and vegetation can be seen in the reflectance spectrum of vegetation shown in Figure 2.1. It should also be noted, that as the leaves mature, more air spaces are formed between the cells in the leaves, causing greater reflection in the NIR wavelengths. Additionally, as canopies of leaves become more dense, the added layers of leaves provide additional opportunities for NIR energy to be reflected. Thus, the denser the leaf canopy, the greater the NIR energy reflection.

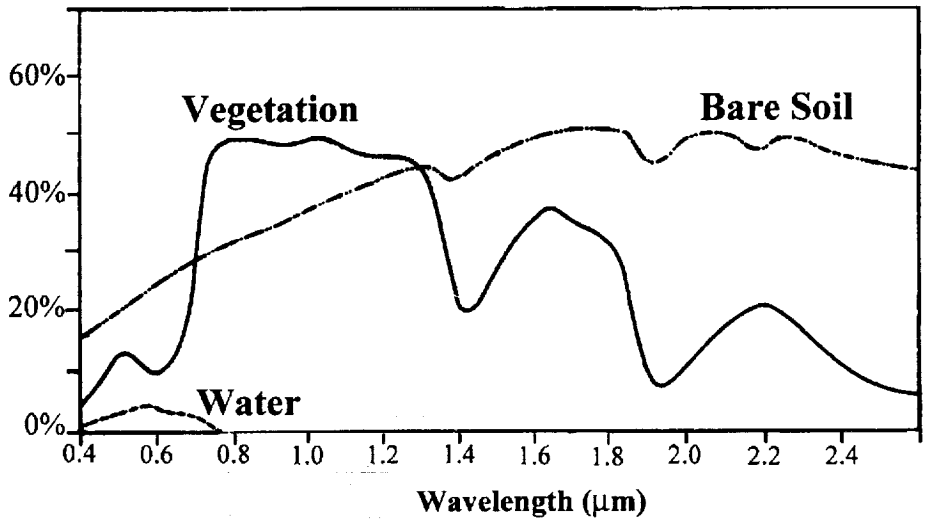



Figure 2.1: Examples of Reflectance Spectra for Vegetation, Soil, and Water, adapted from [21]

## 2.3 Vegetation Indices

Considerable research has been conducted on the development and testing of vegetative indices which, generally speaking, are transformations or combinations of original spectral reflectances collected by remote sensors. Many indices have been developed which are sensor-specific, requiring specific bands in the visible and infrared portions of the spectrum. A widely used index that is applicable

to any sensor with a channel in the visible portion of the spectrum and one in the near infrared portion, is called the Normalized Difference Vegetation Index (NDVI). Through simultaneous collection of overhead imagery and ground measurements, NDVI values have been found to be correlated with vegetative canopy characteristics such as leaf area index [28]. Numerous other studies of the use of NDVI, and other vegetation indices, derived from a variety of sensors, have been conducted. Among them are [32], [33], [36], [11], and [22]. Investigations of the use of vegetation indices to study specifically rice were conducted by [2], [20], [8], [37], [38], and [40].

The NDVI is calculated using the reflectance in the visible (noted as VIS) and NIR channels for each pixel in the image using the following formula:

$$NDVI = \frac{Reflectance_{NIR} - Reflectance_{VIS}}{Reflectance_{NIR} + Reflectance_{VIS}}$$

The resulting NDVI values range between -1 and 1. Because, as described in the previous section, vegetation has higher reflectances in the NIR wavelengths than in the visible wavelengths, NDVI values for vegetation are always greater than 0, and as the plants mature and canopies become more dense, the NDVI values can be generally expected to increase through the growing season. As vegetation starts to die off, or become stressed due to disease, the chlorophyll levels in the leaves begin to decrease and subsequently, the absorption of visible wavelength energy will decrease as well. Since in this case, the reflectances collected by the visible channels in a sensor start to increase, the NDVI values will begin to decrease. An additional use of NDVI derives from the fact that water, soil, and man-made objects do not have the spectral response functions at all similar to vegetation. Thus the NDVI values for these objects are lower than for vegetation. Thus, NDVI is very often used to differentiate vegetated areas from non-vegetated areas.

Other vegetation indicies have been developed and used over the years,

especially those which are capable of correcting for contributions of reflected energy from the backgrounds of vegetation. An example of one of these is the Soil Adjusted Vegetation Index (SAVI). As the name implies, SAVI adjusts the reflectances collected by a sensor such that effect of soil reflectance is removed, leaving the index to record the response due to vegetation only. The use of SAVI requires that the soil background of the vegetation of interest be isolated by collecting a raw soil reflectance. In the case of rice paddies, the background for a majority of the growing season is water, and not soil. Because the paddies are flooded only after the rice plants have begun to grow, it is not possible to obtain a reflectance value for the water only. It was thus decided that the only vegetation index that would be investigated in this thesis was NDVI.

## 2.4 Application of AVHRR to the Study of Rice

The subsection below will describe the attributes of the AVHRR sensor which can be utilized to provide useful information on vegetation. Following that will be a description of the manner in which the imagery was processed and prepared for application in this study.

### 2.4.1 The AVHRR Sensor

The AVHRR sensor is carried on-board the National Oceanic and Atmospheric Administration's (NOAA's) Polar Orbiting Environmental Satellites, which are sun synchronous platforms. The sensor acquires data in five channels, which are defined in Table 2.1 for the most recent platforms. The spatial resolution of the sensor is 1.1 km at nadir and approximately 6 km at the edge of the scans (55.4 degrees off-nadir). The relatively coarse spatial resolution allows for images of extensive land areas to be collected, from which studies of large-scale vegetation

Table 2.1: AVHRR Channel Specifications, adapted from [13]

| Channel Number | Description | Band Width |
|---|---|---|
| 1 | visible | 0.58 to 0.68 $\mu m$ |
| 2 | reflected infrared | 0.725 to 1.05 $\mu m$ |
| 3 | hybrid infrared | 3.55 to 3.92 $\mu m$ |
| 4 | thermal infrared | 10.3 to 11.3 $\mu m$ |
| 5 | thermal infrared | 11.5 to 12.5 $\mu m$ |

development cycles, vegetation transformations such as tropical deforestation, and land-cover type classifications can be conducted [27]. One feature of the AVHRR sensor which offsets the low spatial resolution in comparison to other sensors, such as the Landsat Thematic Mapper (TM), is the fine temporal resolution it achieves. As opposed to Landsat TM, which has repeat coverage every 16 days, a given portion of the earth can be imaged between 1-3 times per day with AVHRR, allowing for the ability to monitor vegetation dynamics over shorter time periods. In this study, only the image from the pass which occurred over the area of interest in the afternoon each day was utilized. This pass was selected since it is the closest to solar maximum.

In deciding to utilize remotely sensed data, researchers must often make a decision between the amount of ground area covered in a given image scene and the spatial resolution of the pixels in the scene. Generally speaking, the higher the spatial resolution of an image, the less ground area that image will contain. For studies being conducted on earth systems of regional and/or global scales, scientists quite often have to give up high spatial resolution in order get imagery which covers more of the earth's surface. Because of this, imagery from AVHRR is routinely utilized to conduct studies on regional and global scales.

### 2.4.2 Processing of AVHRR Imagery

The source of the AVHRR data used in this study varied based on year, but all of it was received originally in raw form. Image data from 1992-5 was offloaded from archive tapes held at the Colorado Center for Astrodynamics Research (CCAR). The data on these tapes were backups of transmissions received by a high resolution picture transmission (HRPT) antenna operated by CCAR, which receives direct data transmission from the NOAA polar orbiters. Imagery from 1998 was obtained directly from the HRPT antenna, and the 1999 imagery was received from the University of Texas which operates its own HRPT antenna. Additionally, image sets were augmented when needed by imagery downloaded from the Satellite Active Archive (SAA) website operated by NOAA.

The raw data was then processed in the following manner. First it was navigated using code written at CCAR which uses ephemeris data from the satellite. The next step was to calibrate the navigated data, the values for which were obtained directly from NOAA. Because calibration allows reflectances and temperatures represented by imagery from different sensors to be directly compared, great care was taken to ensure that the most up-to-date calibration coefficients were used. In a case where it was found that the coefficients had been updated, the imagery was reprocessed. After calibration, and before any further processing was done, images which were found to have pixels representing the study areas which exceeded solar and sensor viewing angles of 30 degrees were eliminated.

The next step in the image processing chain was to correct for the effects that the atmosphere has on the reflectances collected by the sensor. These corrections are necessary due to the fact that atmospheric particles and gases can scatter, reflect and absorb electromagnetic radiation such that the reflectances collected by a satellite sensor do not accurately represent the energy originating from the

ground area represented by a pixel. In the wavelength regions being studied in this thesis, namely Channels 1 and 2 of AVHRR, one must be concerned most with scattering of energy in the visible wavelengths, and absorption of energy in the NIR wavelengths. Scattering will generally cause reflectances collected by Channel 1 to be higher than they should be, and the reflectances collected by Channel 2 to be lower, resulting in lower NDVI values than would be with no atmospheric effects.

The algorithm selected in this study to correct for atmospheric effects is a semi-empirical algorithm called: Simplified Method for Atmospheric Correction (SMAC) [29]. One non-image derived input which is required for running SMAC is aerosol optical depth. Because this value is difficult to obtain, an approximation was used in this application which utilizes visibility data [39]. These were obtained from the National Climatic Data Center (NCDC) of NOAA. Other non-image derived inputs required by SMAC are water vapor content and the properties of the aerosols found in the air. The inputs used in this study represent expected values for atmospheric conditions in the mid-latitudes during summertime, and were obtained from a radiative transfer atmospheric correction algorithm called Second Simulation of Satellite Signal in the Solar Spectrum (6S).

Once atmospheric correction was complete, the images were run through a cloud detection algorithm which utilizes a split channel technique. In this study, pixels whose channel 1 8-bit reflectance values were greater than 11% and whose 11-bit channel 4 values were less than 300 degrees Kelvin were considered clouds and masked to a value of zero.

Once each of the individual five-band images were navigated, calibrated and atmospherically corrected, the NDVIs were computed. The NDVI images were then overlaid with a map and nudged, if required, in order to correct for any

satellite time errors that would cause each image to not register to the map, and each other. Then, the final step in the image processing chain was the formation of an NDVI time series for each growing season. The method and rationale for assembling these time series are discussed next.

### 2.4.3 NDVI Time Series

As pointed out in Section 2.4.1, one of the attributes of the AVHRR sensor platform is a high repeat sampling. This characteristic can be taken advantage of, especially when studying vegetation. If one builds up a time series of NDVI images, then one can monitor the changes in the vegetation of interest over time. In this research, a time series "cube" of 10-day composite NDVI maps was assembled for each growing season under study.

Because of the elimination of some images due to solar and sensor acquisition angles, and some pixels due to the presence of clouds, the sampling rate of images over a growing season is not regular. As a means of eliminating this inconsistency, as well as mitigating the effect of cloud-affected pixels which are not detected by cloud algorithms, a common method of assembling NDVI time series is to create a series of maximum value composites (MVCs). An MVC can be assembled for every week, every ten days, every month, etc. Selection of the periodicity of the composites depends on several factors, such as the duration of the overall time series. A pseudo-decadal time period was utilized in this study such that each month had three composites. The first two composites had ten days each, while the third composite had ten or eleven days depending on the total number of days that month has.

The rationale for using the maximum-value for a given pixel over the composite period, rather than, say, an average value, is to maximize the chances of

eliminating the effects of clouds, since the NDVI of a cloud is lower than for vegetation. Another advantage in using the MVC technique is that the effects of the anisotropic nature of reflectances collected by satellites are minimized. Because a reflectance measurement is particular to its set of acquisition and illumination geometries, it would be ideal to take into account the bidirectional reflectance function (BRDF) for each material type in the image. In this manner, corrections could be made to allow a more sound comparison among images collected of the same area at different times [46]. Unfortunately, to implement a BRDF correction on large images covering many years worth of data is, for many reasons, an intractable goal. The impact of the inability to accomplish this task can be mitigated by using MVCs because the geometries which result in higher NDVI values are favored, creating a higher level of consistency from image to image. Thus, a time series resulting from the use of MVC NDVI values may be slightly higher, on average, than the time series would have without MVCs, but this minimizes the chances that variations in NDVI value within a time series are due to only BRDF variations.

A schematic drawing representing the concept of the AVHRR NDVI time series that were built for each of the years under study in this research is presented in Figure 2.2. These image "cubes" allow one to look at a single pixel, representing a particular area of interest, through an entire growing season, and this concept forms the basis for all further analysis in this research. These time series are used to define the growth of rice plants, the purpose of which is to couple the image data with methane flux estimation models so that there is less reliance on ground truth data. These time series image cubes are also used in an effort to detect rice growing areas. These topics will each be addressed in the following chapters.

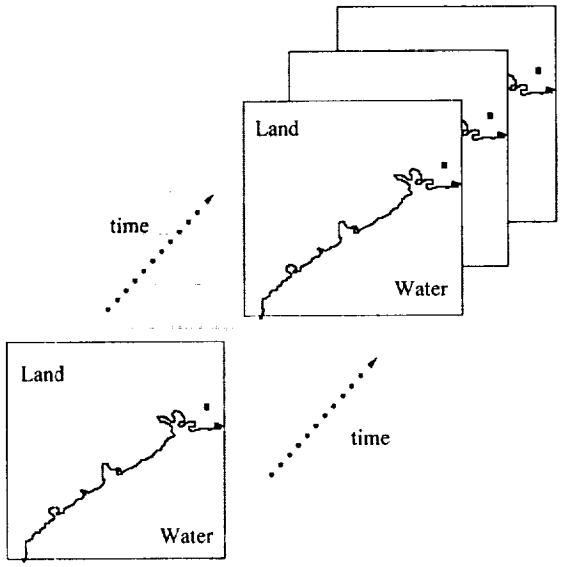



Figure 2.2: Schematic of NDVI Time Series Cube

# Chapter 3

# Estimation of Rice Biomass Using AVHRR

## 3.1 Introduction

As mentioned in the Chapter 1, one of the byproducts of rice cultivation is methane, a gas whose increased levels in the atmosphere has garnered greater attention recently. The overall goal of this research project is to explore the relationship between AVHRR imagery and rice production, and to ascertain the extent to which the data derived from AVHRR can be successfully coupled with methane flux models. In order to obtain direct input to the models from the satellite imagery, a relationship must be found between the NDVI time series, and total aboveground biomass. This chapter describes the data that was used to derive this relationship, the methodology used, and the results.

## 3.2 Rice Plant Ground Truth Sources

An excellent source of plant data for this study was provided by Drs. Gowei Yu and Ted Wilson of Texas A&M University (TAMU). The data was collected for the 1992 - 5 growing seasons at the TAMU Rice Research Station, located near Beaumont Texas, and for the 1998 - 9 growing seasons at a farmer's research field near El Campo Texas. A reference map is provided in Figure 3.1.

The plant data collected over the six years includes almost a dozen different

rice varieties that are commonly grown in Texas, and which represent a wide range of growth characteristics. Included in the data that were collected were LAI and total aboveground biomass measurements (for the remainder of this report, the phrase "total aboveground biomass" will be referred to simply as biomass). The sampling methodology utilized by Wu and Wilson was a complete randomized plot design, with each variety having between 9 - 12 plots (depending on the year) for destructive sampling, and 9 - 12 plots for yield estimation. Sampling frequency was roughly twice a week at the beginning of the growing season, and roughly once a week thereafter. A subplot was randomly selected for each destructive sampling, making sure that each of the sample areas was separated by at least 0.5 meter of undisturbed plants [49].

Because the LAI and biomass data were collected with non-regular time intervals, the data were sampled for use in this project to match the 10-day periods of the NDVI time series. If there were multiple samples in a given 10-day period, then the values were averaged. There were some isolated cases where there were 10-day periods for a given variety that had no data recorded. In those cases, data points were interpolated based on the percent change measured for other varieties during the same 10-day period.

## 3.3 Derived Biomass Estimates

Beginning with the 6 sets of annual NDVI values, and LAI and biomass measurements, a means of relating any or all of the data was sought. Linear regressions were calculated between NDVI and LAI, and NDVI and biomass, with some satisfactory results. But it was important to derive a relationship that while strictly empirical, made some physical sense as well. As described in Chapter 2, NDVI is a good indicator of the amount of vegetation and the vigor of the

vegetation. More vegetation, and more vigorous/healthy vegetation is directly related to the photosynthetic potential of the plant. It is the photosynthetic process that is responsible for the production of plant mass. Because biomass during a growing season is essentially an accumulation process, it was reasoned that comparing the accumulating NDVI over the growing season should provide a reasonable proxy for the time series of biomass. Further support for selecting to estimate biomass rather than LAI is provided by Casanova et al. who state that "biomass during the rice-growing cycle is more precisely estimated [using reflectance data] than LAI" [8]. The plot shown in Figure 3.2 is the cumulative NDVI time series plotted against the biomass data obtained from the two ground truth sites for each of the six years. Because during the first twenty days of the growing season the plants are too small to be driving the NDVI of the pixels, it was decided to not include the first two 10-day periods in any further analysis.

To get a better sense of the validity of the reasoning given above for relating cumulative NDVI with biomass, NDVI was plotted against the changes in biomass from each 10-day period to the next - in essence, un-accumulating the biomass. For simplicity in the plot shown in Figure 3.3, the averages over the six years are shown. Except for the 5th and 10th 10-day periods, the trends of the two time series track each other, providing some added credibility to the supposition that cumulative NDVI and biomass are related physically.

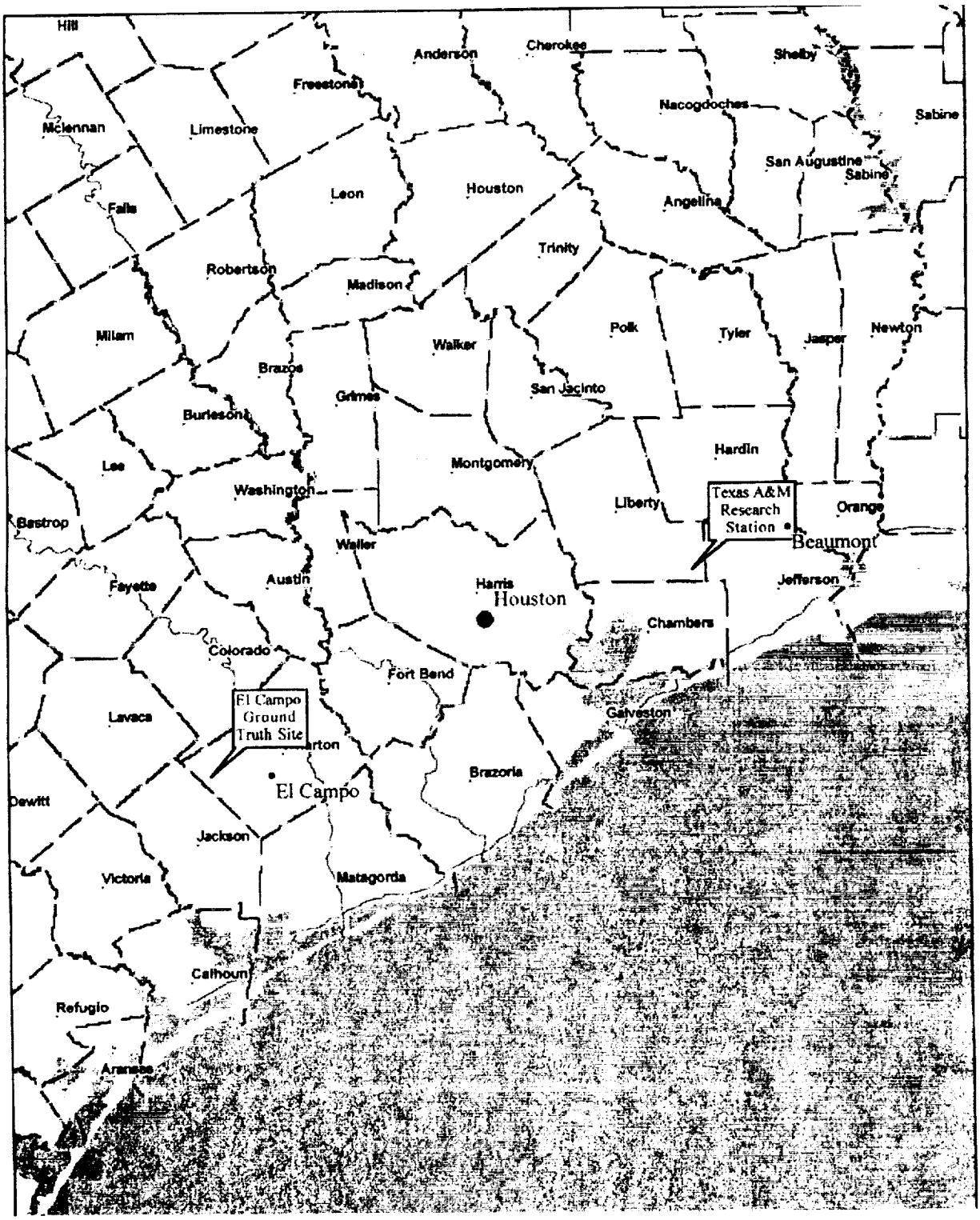



Figure 3.1: Rice Growing Region of Texas with Notations of Ground Truth Sites

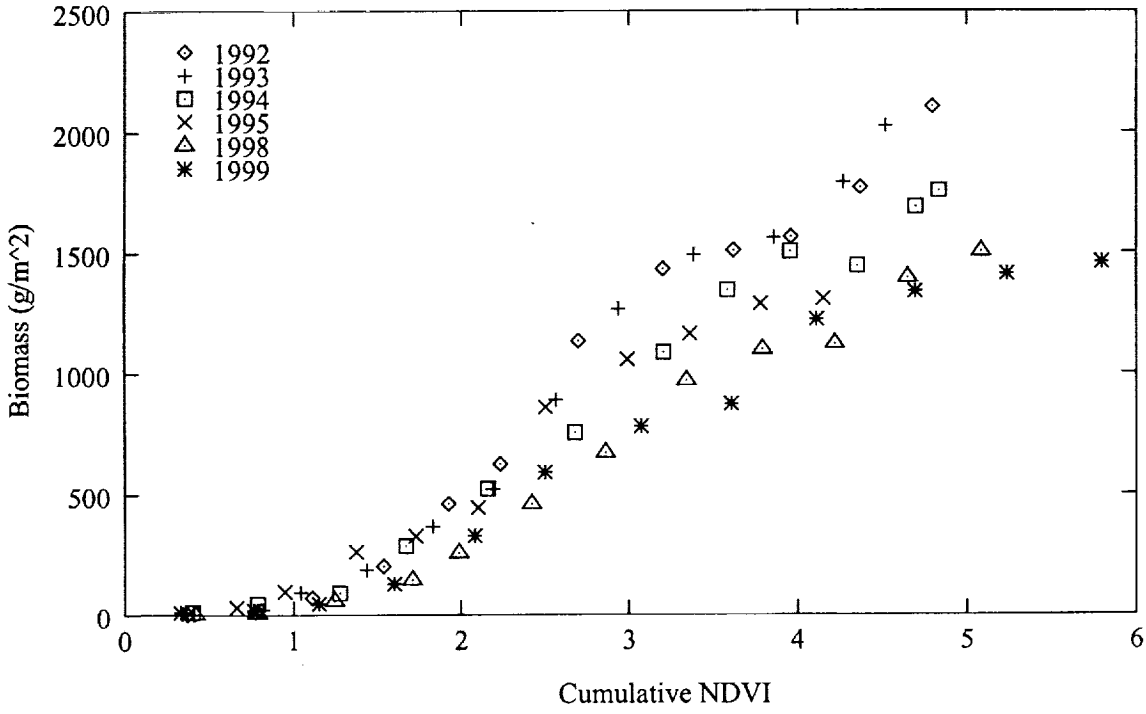



Figure 3.2: Cumulative NDVI vs. Total Aboveground Biomass for 1992-5 and 1998-9

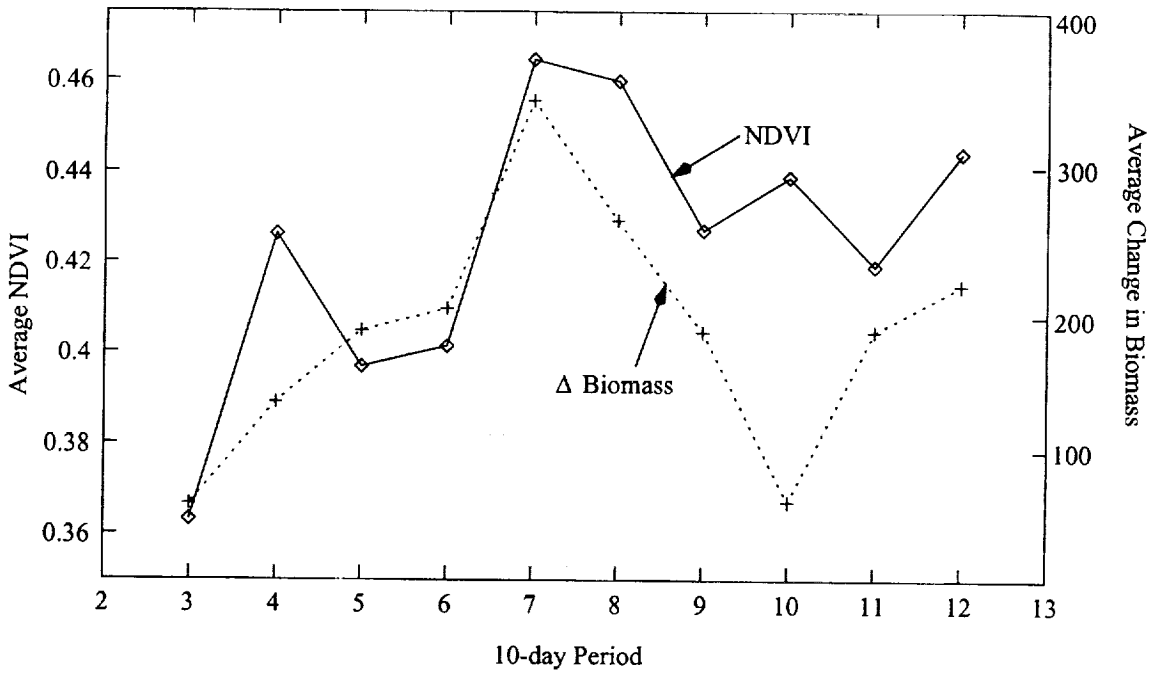
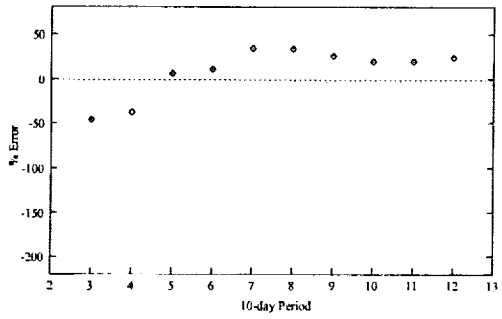


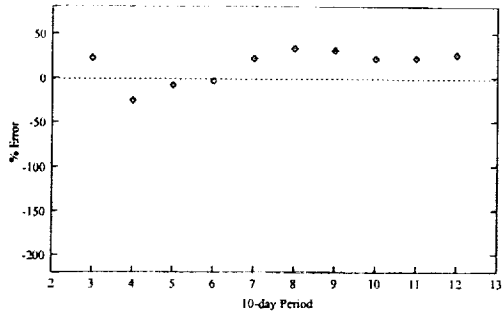

Figure 3.3: Average NDVI and Average Delta Biomass Values over 1992-5 & 1998-9

As a means of assessing the general magnitude of errors in the estimation of biomass that may be expected with this method, the following analysis was done. First, regression relationships were calculated for the six possible combinations of five years. Then, the biomass for the year that was left out of the formation of the regression relationship was estimated from the cumulative NDVI for that year. Finally, the estimated biomass was compared to the measured biomass. For example, all of the cumulative NDVI and biomass data were pooled for 1993-1995 and 1998-1999, omitting the data pairs for 1992, and a regression relationship was calculated. Then, the 1992 biomass was estimated by evaluating the cumulative NDVI values for 1992 through the regression relationship. This was done for each of the six years under study. The results of the estimation error are shown in Figure 3.4, where %error was calculated as: ((Estimated-Measured)/Measured)*100.

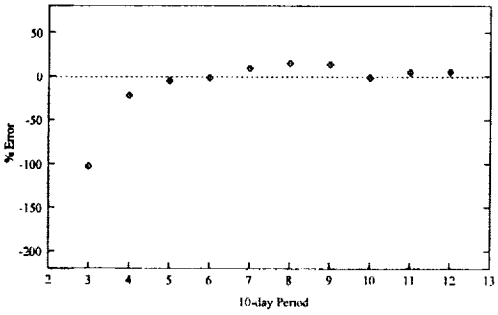
The errors in estimating 1992-1995 biomass are quite reasonable, and for most of the years, the average error is driven by the first couple of biomass es-



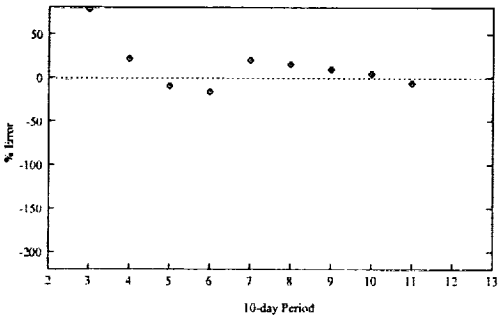

(a) 1992: Average Error = 9%



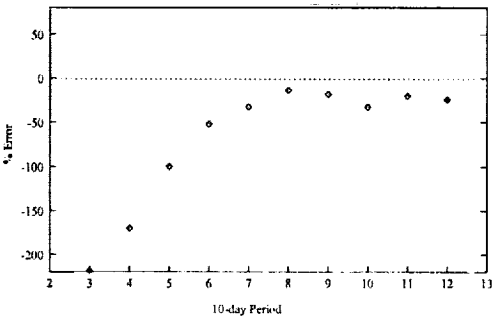

(b) 1993: Average Error = 14%



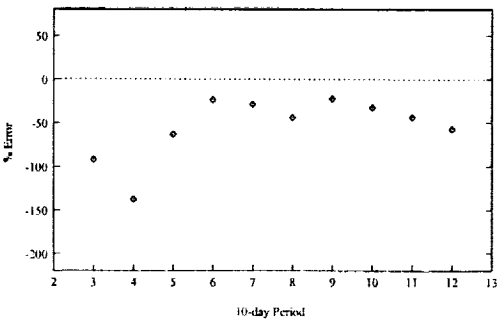

(c) 1994: Average Error = -9%




(d) 1995: Average Error = 13%




(e) 1998: Average Error = -68%




(f) 1999: Average Error = -54%

Figure 3.4: Percent Errors in Calculating Biomass Using Cumulative NDVI-to-Biomass Relationships

timates in the 10-day period. The errors for 1998 and 1999 are, however, more problematic. As can be noted from the plot presented in Figure 3.2, the 1998 and 1999 data tends to have higher cumulative NDVI values coupled with lower biomass values than do the 1992-1995 data. In an effort to try to understand the source of the difference, it was first verified that the rice varieties that were planted in 1998 and 1999 were also used in some of the data collected for the earlier years. Additionally, no systematic differences were found in weather, such as temperature and rainfall, between the two groupings. The last known variable different between El Campo (1998-1999) and Beaumont (1992-1995) is soil. As one can note from the information provided in Table 3.1 the soil components for the two areas are quite different.

Table 3.1: Composition of Soils at El Campo and Beaumont Ground Truth Sites

| Constituent | El Campo | Beaumont |
|---|---|---|
| Sand | 63 % | 25 % |
| Clay | 13 % | 32 % |
| Silt | 24 % | 43 % |

Despite the noted differences in the El Campo and Beaumont NDVI-biomass relationships, it was decided to continue with the regression relationship formed with all six years-worth of data since the goal of this research is to derive a method which may be useful for entire regions. This final regression equation, shown in Figure 3.5, provides the means of estimating biomass values for the rest of the study.

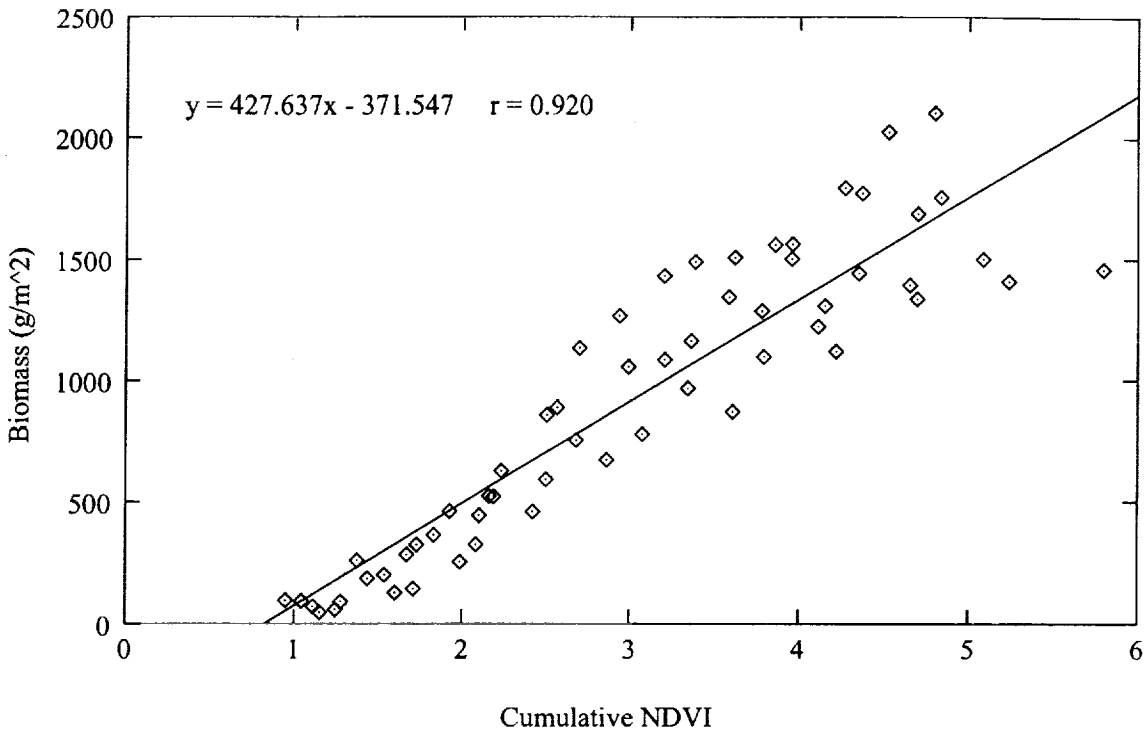
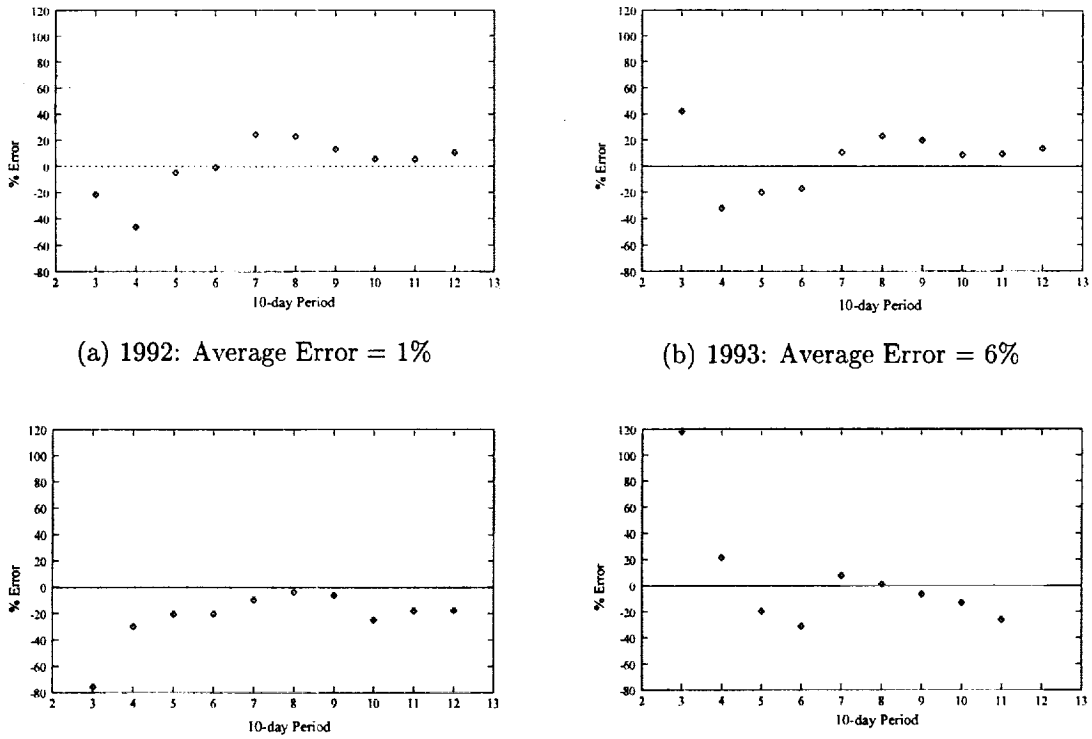



Figure 3.5: Regression Relationship Derived for Cumulative NDVI and Biomass for 1992-5 & 1998-9

However, in order to further look at the effect of including seemingly disparate regions in a relationship between cumulative NDVI and biomass, the results from a calculation of a different regression relationship for El Campo and Beaumont are now presented. Similar to the analysis related to Figure 3.4 presented above, the first step was to calculate regression relationships for each permutation of 3 of the 4 years of data collected at Beaumont. In each case, the year that was left out of the relationship was used to test that relationship by comparing the estimated biomass with the measured biomass. The results for each are presented in Figure 3.6.




(a) 1992: Average Error = 1%

(b) 1993: Average Error = 6%

(c) 1994: Average Error = -22%

(d) 1995: Average Error = 6%

Figure 3.6: Percent Errors in Calculating Biomass Using Cumulative NDVI-to-Biomass Relationships (1998 and 1999 Removed from Analysis)

As can be noted in Figure 3.6, the average error for the predictions, using regressions formed with only the Beaumont data, decreased for 1992, 1993, and 1995, while it increased significantly for 1994. In light of this result, it seems that utilizing a separate regression relationship for the two different locations may be advantageous. However, the dramatic increase in error for 1994, and the fact that this method of guaging error levels could not be performed for the El Campo site (due to the fact that there were only two years of data collected), it was decided to proceed with the rest of the analysis using the 6-year regression shown in Figure 3.5. For reference, the two regression relationships are provided below in Figure 3.7.

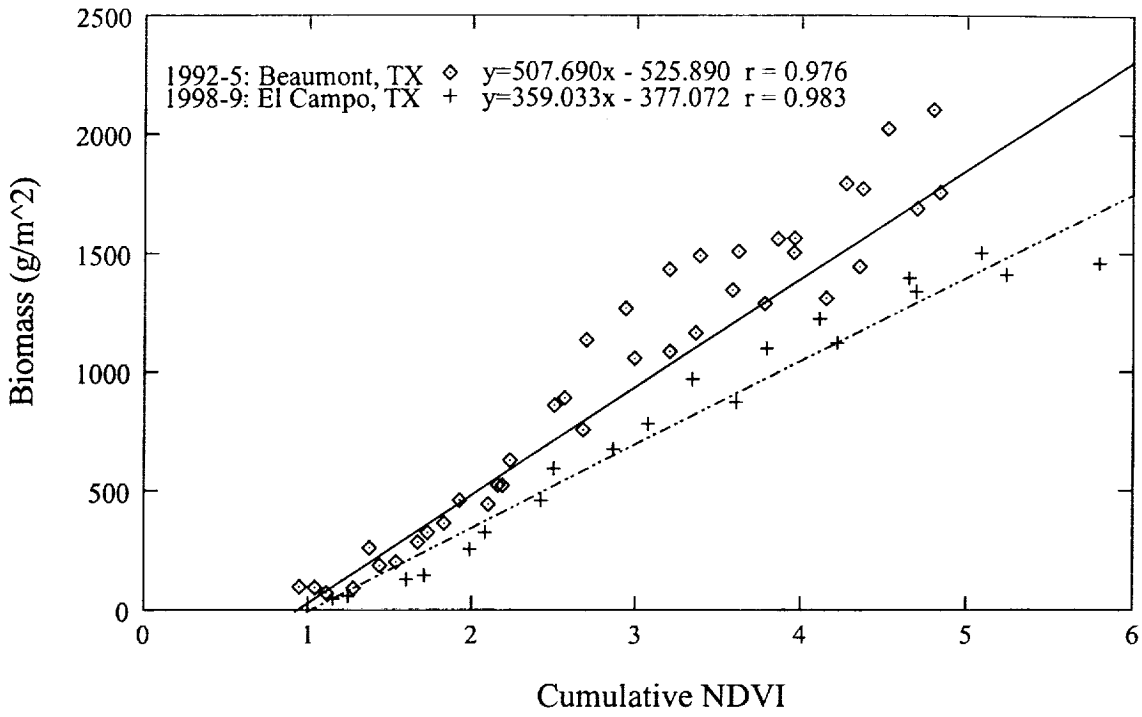



Figure 3.7: Split Regression Relationships Derived for Cumulative NDVI and Biomass for 1992-5 & 1998-9

# Chapter 4

# Coupling AVHRR-Derived Biomass with a Semi-Empirical Methane Emission Model

## 4.1 Introduction

The semi-empirical model discussed in this chapter was developed by Yao Huang as a part of doctoral research at Rice University in Houston Texas, under the supervision of Drs. Ron Sass and Frank Fisher. The data used in deriving the relationships incorporated into the model were collected during field experiments at the Texas A&M Beaumont Research Station and in pot experiments run on the Rice University Campus. The model was validated using some Beaumont field data which were held back from the development data, as well as on data for several Asian sites. The following is a description of the Huang model, including the processes it takes into account and how they are modeled, the manner in which the remote sensing data was coupled with it, and the results that were obtained.

## 4.2 Methane Flux Ground Truth

The two methane flux estimation models, as will be described in this chapter and the next, require biomass inputs which until now, could only be derived from direct measurements. These requirements make application of these models to locations other than research areas where such data is collected almost intractable.

A method for utilizing AVHRR NDVI time series to estimate biomass was presented above. However, in order to assess the success of the method, comparisons must be made between the flux estimates made by the models (both with and without the remote sensing data inputs) and real flux data. In order to do this, methane flux measurements collected during a time for which AVHRR imagery are available must be utilized.

Methane flux ground truth data was obtained from Drs. Yu Huang and Ron Sass of Rice University who collected the data at the Texas A&M Rice Research Station near Beaumont in 1994. The methane flux measurements were collected from paddies different from the ones where the biomass observations were collected by Drs. Wu and Wilson. The measurements were collected from four randomly selected measurement sites using a chamber method. The headspace gas of the open-bottom chamber was collected approximately twice a week and five samples were collected for each chamber over a 30 minute period. The gas samples were then analyzed in a gas chromatograph equipped with a flame ionization detector in order to obtain the methane mixing ratios [15].

## 4.3 Huang Model Description

In recognizing the role that the rice plant plays in the production, oxidation, and transport of methane, the Huang model was formulated with an emphasis on the growth and development of the plant. For the methane production cycle, the model considers two sources of substrates for the methanogenic bacteria, the rice plant itself and added organic amendments. Additionally, some environmental factors, such as the percentage of sand in the soils, that affect the production were included. Using data collected in past studies, Huang formed a relationship estimating the amount of carbohydrates derived from the rice plant with above-

ground biomass. Another empirical relationship was formed for the amount of carbon substrates made available by organic matter amendments. The final component in Huang's estimate for the amount of methane that is produced is soil redox potential.

Methane oxidation and transport are not specifically addressed in the Huang model in the form of equations. However, use of data from past studies provided fractional coefficients for the methane production relationships which account for the amount of produced methane which is immediately oxidized after production, and then the amount of the remaining methane which is finally transported.

Again, the emphasis of the Huang model is the plant growth cycle. In the absence of daily biomass measurements to provide to the flux model, Huang simulates biomass accumulation using a logistic growth equation of the form

$$W = \frac{W_{max}}{1 + B_0 exp(-rt)} \tag{4.1}$$

where

$$B_0 = \frac{W_{max} - W_0}{W_0} \tag{4.2}$$

and

$$W_{max} = 9.46GY^{0.75} \tag{4.3}$$

Plant aboveground biomass on a given day is $W(gm^{-2})$. $W_0$ and $W_{max}$ are aboveground biomass at the beginning of permanent flooding and at the end of a growing season, respectively. The variable $t$ is the time scale measured in days after permanent flood. The constant $r$ is an intrinsic growth rate for biomass and was empirically derived.

In summary, the Huang model bases its estimates of daily methane flux from rice paddies based on production levels which are empirically derived relationships.

These relationships require an estimate of the daily biomass which are simulated using the logistic growth equation shown above. Required inputs are crop yield, to calculate the $W_{max}$ value, and $W_0$. However, it was not clear from where the $W_0$ value was expected to come in future applications.

## 4.4 Implementation and Results

The first task in implementing the Huang model was to generate the daily biomass values to use in the model. Information was available to generate the biomass time series using Huang's logistic growth curve equations, with a value for $W_0$ being suggested by Dr. Huang.

Next, a means of generating daily biomass estimates from the AVHRR NDVI-derived 10-day biomass estimates needed to be found. Two methods were employed. The first was to fit a second-order polynomial through the AVHRR-derived biomass estimates, and then to use the polynomial equation to generate the daily values. The second method was to utilize the logistic equation suggested by Huang but to utilize AVHRR derived $W_{max}$ and $W_0$ values.

The three resulting daily biomass curves are shown in Figure 4.1. Also included for reference on the plot are the biomass ground truth data points collected by Huang, as well as the individual biomass estimates derived from the AVHRR-NDVI time series.

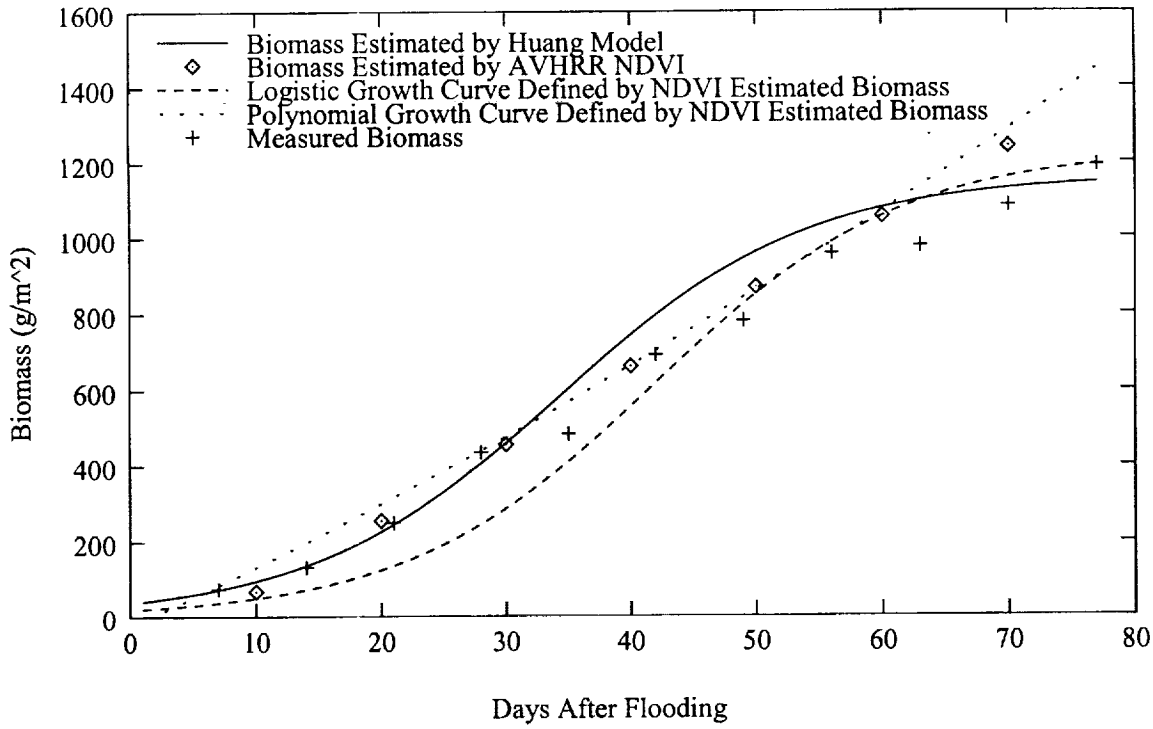



Figure 4.1: Comparison of Measured and Daily Predicted Biomass Values

There are several things to note on this plot. First, observing the filled and open circles, one can note that the AVHRR NDVI-derived biomass seems to provide a good estimate of actual biomass, adding credibility to the cumulative NDVI-to-biomass relationship derived for this part of Texas. Another overall observation that can be made is that all three of the daily biomass curves provide similar results. Looking next at the daily biomass time series derived from the logistic relationship using the Huang inputs, one can note that it does an excellent job of detecting the first thirty days of actual biomass values, but that it begins to overestimate the biomass after that. The logistic growth curve generated by using the biomass estimated by AVHRR NDVI clearly underestimates the measured biomass for the first two-thirds of the growing season, after which it begins to overestimate. So, on average, the NDVI-derived daily biomass curve seems to fit the measured data better. However, because methane flux from rice plants is not

uniform over the growing season, having a better on average fit is not necessarily better.

Once the daily biomass curves were generated, each was run through the flux estimation model. In addition to the biomass values, several other inputs were required. The first is an average soil temperature for the growing season. This was derived from the average air temperature using the relationship $T_{soil} = 4.4 + 0.76T_{air}$ provided by Huang. Daily air temperatures were acquired from NCDC and then averaged, as required by the model. Another required input is a soil index, $SI$, which is based on the percentage of sand in the soil using the relationship $SI = 0.325 + 0.0225(\%sand)$. This index is an empirically derived relationship which is intended to characterize the relative effect of soil texture on methane production and emission. Finally, a vegetation index, $VI$, is required. This scaling factor too is based on an empirical relationship [16] and is intended to account for relative differences in methane production among rice varieties. Texas rice varieties are generally assigned a $VI$ of 1 or 1.5 [15]. Recognizing that several varieties of rice are likely found in the area represented by a given AVHRR pixel, a value of 1.25 was used for this study. The results of using the three daily biomass series are shown in Figure 4.2. Also included on the plot are the measured flux values.

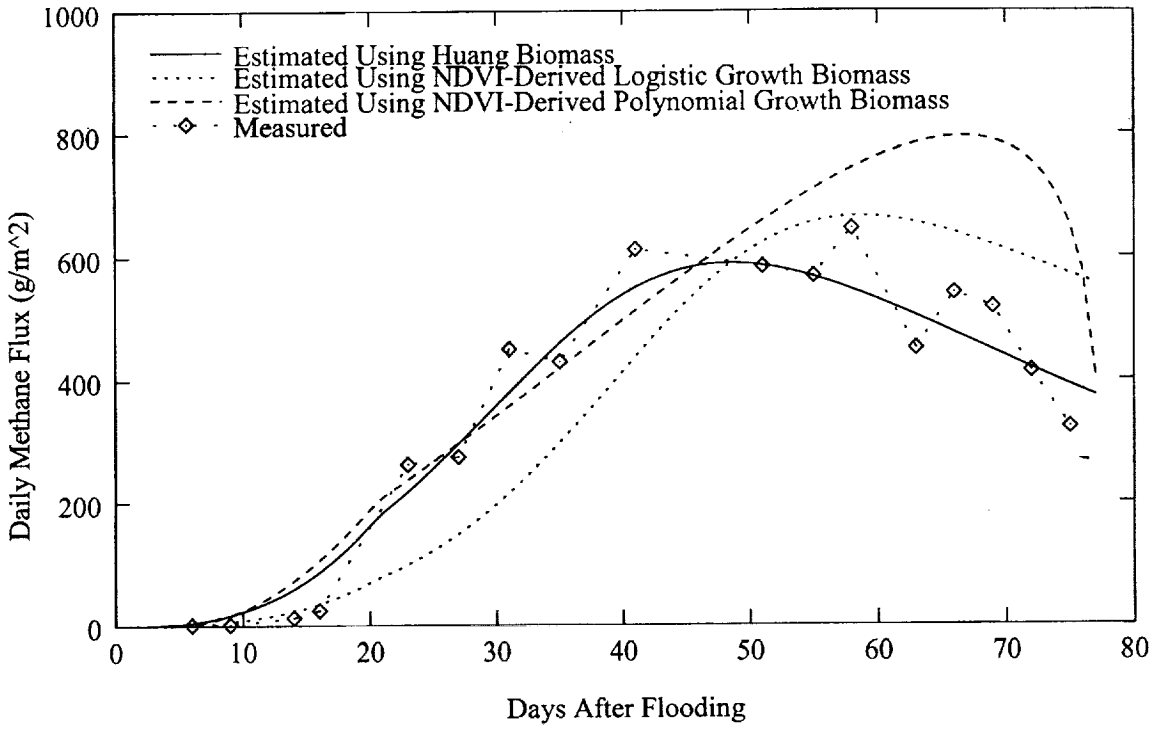



Figure 4.2: Comparison of Measured Methane Fluxes and Predictions from Huang Model

Looking first at the daily methane flux curve that resulted from the AVHRR-derived logistic growth biomass data, one can note that it significantly underestimates the flux. This is not surprising, because the method of generating the growth curve resulted in biomass values that underpredicted the measured values. Because methane production is dependent upon the availability of substrates, and because those substrates result from plant growth, predictions of biomass which are lower than reality, will subsequently result in methane flux underestimation. Another thing to note with the AVHRR-derived logistic growth biomass input is that the flux estimates do not decrease at the same rate at the end of the growing season as the measured values. This can be traced back to the overestimation of the biomass with a logic similar to the cause of the underestimation of the fluxes earlier in the season.

Looking next at the daily methane flux curve which resulted from the Huang-

derived biomass, we note a good fit with the measured flux values throughout the season. The only criticism with this estimation is that it misses the methane flux peak at about day sixty and thus begins its decrease too soon.

The daily methane flux curve resulting from the AVHRR-derived polynomial growth biomass provides an excellent fit to the early season flux measurements, however, the estimates during the last twenty days of the season are extremely high. This is due to the fact that the daily biomass estimates are significantly higher at the end of the season than the measured values. The polynomial which resulted from the curve-fitting of the 10-day estimated biomass was essentially linear, with the coefficient of the second order term having very low value. This suggests that indeed, a logistic equation is more appropriate than a quasi-linear relationship of biomass over time. From a purely physical standpoint, this is very logical since plants do not continue to grow infinitely.

Another way to analyze the appropriateness of coupling the Huang model with remote sensing data is to look at total seasonal methane flux rather than daily flux. For atmospheric scientists, daily methane flux estimates are not as useful as total seasonal estimates. To do this, the daily methane flux estimates from the Huang model were simply accumulated to get a seasonal value. For the measured methane fluxes, in order to obtain a figure for the total seasonal flux, the daily values were averaged, and then this average was multiplied by the length of the growing season. This value could then be compared with the summed daily flux values for the three biomass time series methods. Also evaluated for comparison, was the result of a simplified set of equations provided in the Huang model, which utilize only single values to represent the growing season, rather than performing the calculations on a daily basis. In this calculation, the average biomass value is defined to be 55% of $W_{max}$.

As can be noted from the data presented in Table 4.1, both the Huang daily biomass curve and the AVHRR-derived logistic biomass curve provide excellent estimates of both total seasonal and average daily methane flux.

Table 4.1: Summary Comparison of Estimated and Measured $CH_4$ Fluxes for Huang Model

| Biomass Curve | Model Version | Total Seasonal $CH_4$ $(mg/m^2)$ | Average Daily $CH_4$ $(mg/m^2)$ | Percent Error |
|---|---|---|---|---|
| Huang (Logistic) | Full | 27,281 | 354.3 | -2% |
| | Simple | 26,536 | 344.6 | -4% |
| AVHRR (Logistic) | Full | 27,123 | 352.3 | -2% |
| | Simple | 28,896 | 375.3 | 4% |
| AVHRR (Polynomial) | Full | 33,510 | 435.2 | 21% |
| | Simple | 36,067 | 468.4 | 30% |
| **Measured** | | **27,720** | **360.0** | |

Additionally, flux estimates from the simplified version of the Huang model are only slightly worse than those obtained from the full model. Clearly, despite the fact that the plot in Figure 4.2 showed that the polynomial-fit biomass series obtained from the AVHRR-derived biomass estimates had such an excellent fit to the measured fluxes, the overestimation at the end of the season resulted in estimation errors of, at best, 21%.

The sensitivity of the Huang model to differences in biomass were demonstrated through interpretation of Figure 4.2. In an effort to understand better the impact of uncertainty in one of the other primary inputs to the model, soil sand, the percentage of sand in the soil was varied around the measured value for the Beaumont site. The effect of these changes on daily methane flux can be seen in Figure 4.3.

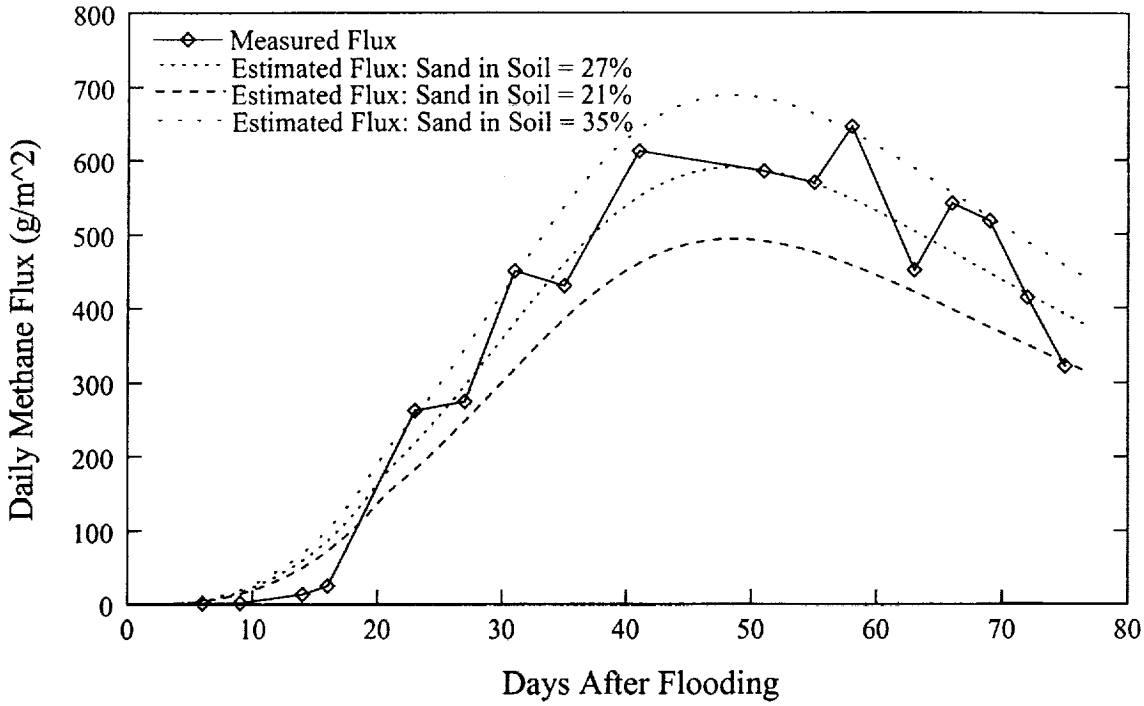



Figure 4.3: Sensitivity of Huang Model to Soil Sand Content

The effect on total seasonal methane is as follows: 27.3 ($g/m^2$) with the measured sand of 28%, 22.8 ($g/m^2$) for soil sand content of 21%, and 31.9 ($g/m^2$) for 35% soil sand. This equates to a difference of 16% change in total seasonal methane flux estimates for a 25% error in soil sand content. Clearly, accuracy in soil sand content is very important in using the Haung model to estimate methane flux.

Soil sand content is not something which is able to be estimated using remote sensing. However, for this model, using biomass estimated from AVHRR imagery can replace often inaccurate yield inputs to drive the biomass input. As an illustration of how inaccuracies in yield can affect methane flux estimates, yield values for the state of Texas and for Jefferson county, in which the Beaumont site resides, were substituted for the yield data collected at the test site. The results of this are shown in Figure 4.4.

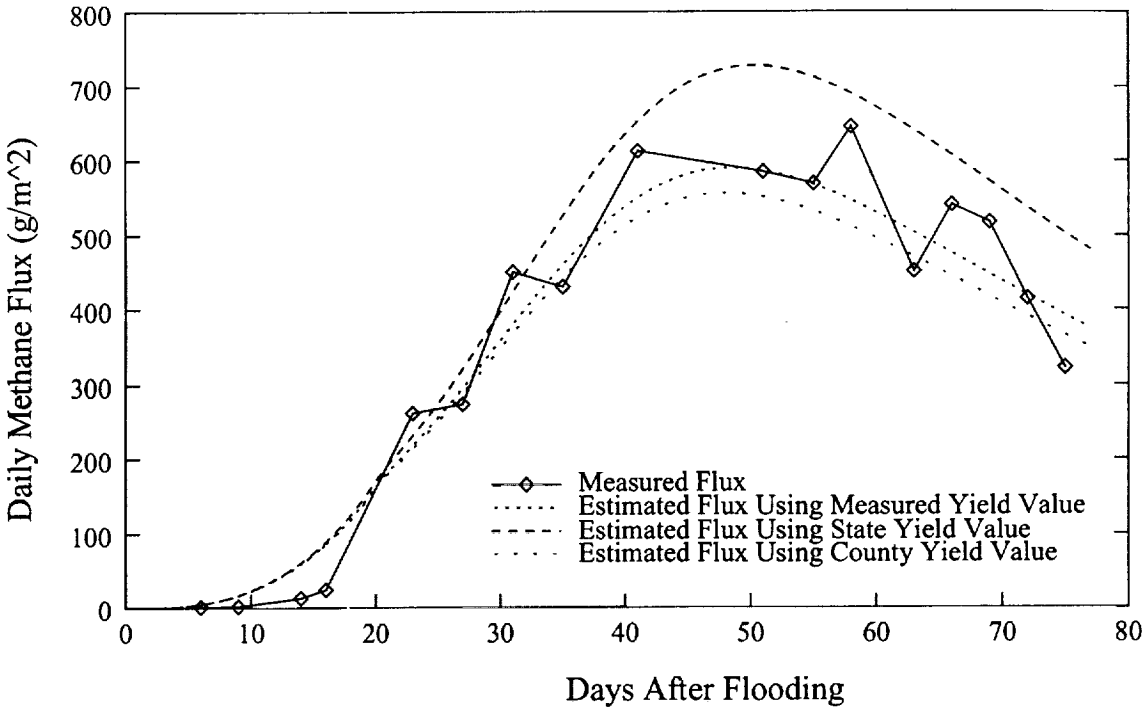



Figure 4.4: Sensitivity of Huang Model to Yield Inputs

These daily flux differences translate into total seasonal flux of 33.0 ($g/m^2$) of methane using the state yield value, compared with 27.3 ($g/m^2$) of methane using the actual yield value for the test site. Using the county yield value netted a total seasonal methane flux of 25.8 ($g/m^2$). It is quite evident that using AVHRR derived biomass yields better results than do inaccurate yield data.

A summary of the results from the coupling of AVHRR-derived data with the Huang model will be presented in Chapter 7, as will suggestions for further research.

# Chapter 5

# Coupling AVHRR-Derived Biomass with a Process-Based Methane Emission Model

This chapter describes the use of AVHRR-derived biomass in a process-based model developed by Peter van Bodegom as a part of doctoral research at Wageningen University in Wageningen, the Netherlands under the supervision of Drs. J. Goudriaan, P. Leffelaar, and A. J. M. Stams. The numerous subcomponents of the model were developed using laboratory data and then validated on several data sets, primarily from Southeast Asia. As will be described more fully, the model was altered specifically for this study to take aboveground biomass as an input. The following sections will describe the model, how it was implemented in this study, and the results of its application to Texas rice crops.

## 5.1 van Bodegom Model Description

The van Bodegom model takes into account more subprocesses which play a role in methane emission from rice paddies and treats these processes in a more rigorous manner than does the Huang model. This is especially true in the production and oxidation cycles. Additionally, the model looks at both of these processes separately for the bulk soil, and the rhizosphere. The rhizosphere is the volume of soil immediately surrounding the root system, and the rationale for treating it separately from the bulk soil is that it is really the effects of the

roots that drive the production and provide most of the means for the oxidation of methane.

In the equations defining methane production, van Bodegom takes into account substrate production from mineralization of carbon available in the soil, decomposition of organic amendments and plant roots, and root exudation. There are terms in these equations which account for the competition that methanogenic bacteria has from nitrite, iron, and sulfate reducing bacteria. Thus, knowledge of the soil beyond the percentage of sand needs to be known. Knowledge of ferric iron, nitrate, sulphate, and carbon contents is required.

In the treatment of methane consumption via oxidation under normal flooded conditions the equations are defined to be related to the growth of the rhizosphere since most of the oxidation occurs there. Additionally, information on flooding conditions over the growing season are an optional input such that any reintroduction of oxygen to the soil can be taken into account.

Methane transport is treated by van Bodegom in a manner similar to Huang, in that each of the potential transport mechanisms is lumped together into a transport coefficient. However, van Bodegom does have different transport coefficients for the rhizosphere and the bulk soil. The one major addition that van Bodegom makes is to take into account the rush of methane that is prompted from the drying of the soil once the paddy is drained, usually 5-10 days prior to harvest. This increased transport would also be reflected if the paddy dries out and this is included in the aforementioned optional input of flooding conditions over the course of the growing season.

## 5.2 Implementation

Originally, the van Bodegom model was developed to account for plant growth by focusing on the development of the root system, which is responsible for almost every aspect of the methane emitted by rice paddies. This root growth was defined by a logistic growth curve which describes the daily root length density (RLD) during the growing season. Similar to the Huang model, the van Bodegom model obtains a value for $RLD_{max}$ through the use of crop yield. Additionally, a harvest index, $HI$, is utilized, which is defined as a ratio of grain yield to total above ground dry mass.

For application to this research, Dr. van Bodegom altered the original model to accept a table of total aboveground biomass values available throughout a growing season. This adaptation now allows the growth of the roots to be dictated by the accumulation of biomass. The biomass values used to drive the model in this study, were the estimates derived from the AVHRR NDVI time series. In the model, the biomass values are then used to drive the root growth calculations mentioned above, with one other piece of data being required. This input is called specific root length (SRL) and it is the ratio of the root length to the dry weight of the root. This is not a well known value for very many varieties of rice, and no reference to the SRL of any of the common Texas varieties could be found. Thus, the default value given in the model was utilized in this study.

Of the other required inputs, some were readily available. The sulfate and nitrate concentrations were set to zero since these compounds are found naturally only in trace levels, and would have non-zero values only when the soil is treated with fertilizers. In the case of Texas, it is known from field treatment logs that the fields were treated only with urea, which do not contribute to the levels of these two compounds. The level of mineralizable carbon in the soil was rather difficult

to obtain, but a source of the information was finally found and the input was set to 1.0%. The iron concentration of 0.2%, soil porosity of 50%, and water content of 25% at field capacity, were all obtained from the same source [7].

## 5.3 Results

The model was run using these inputs, producing daily flux estimates for the 1994 Texas site. Both the estimated biomass from the AVHRR NDVI time series and the measured biomass were run through the model in an effort to ascertain the sensitivity to differences in biomass. This was an unknown prior to this research, since the model was adapted specifically for this study. The differences between the measured and estimated biomass values produced very little difference in daily methane flux. The plot shown in Figure 5.1 presents the model run with the NDVI-derived biomass, along with the measured flux values.

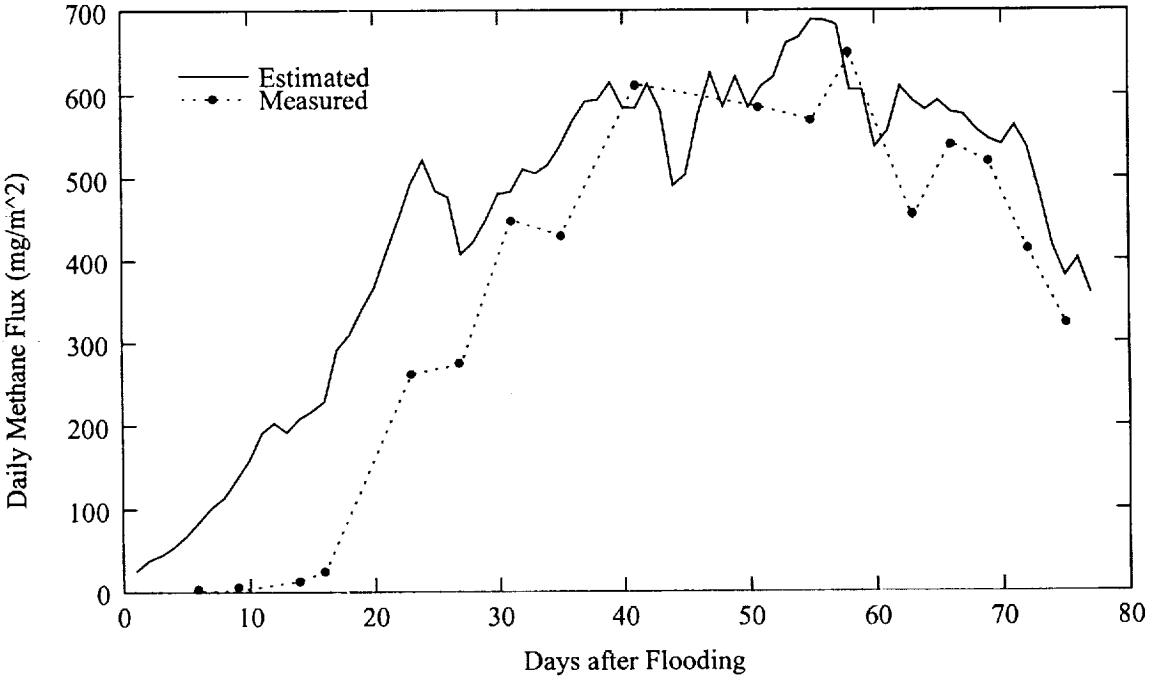



Figure 5.1: Comparison of Measured Methane Fluxes and Predictions from van Bodegom Model

From the plot, it is clear that as currently configured, the van Bodegom model quite accurately follows the trends in the daily methane flux from this Texas field. The only significant deviation from the measured flux is an overprediction for the first 35 days after flooding. This overprediction, however, greatly impacts the average daily and total seasonal flux estimates, which were 448.0($mg/m^2$) and 34,492($mg/m^2$) respectively. These values are almost 25% higher than the measured values of 360.0($mg/m^2$) and 27,720($mg/m^2$). So although the shape of the daily methane flux estimates matches quite closely with the measured flux, it seems that the overestimation at the beginning of the season very much affects the seasonal daily average and total flux. The SRL variable which is unknown for the Texas rice varieties may be affecting this, or more likely, the estimate of soil iron content may not be well enough known.

As a means to understanding better the effect uncertainty in the input variables may have on methane flux predictions, three variables were altered and the resulting predictions were plotted. The three variables selected were biomass, soil mineralizable carbon content, and soil ferric iron content. SRL was not selected for analysis since it is unknown at this time what the magnitude of variation might be. The plot shown in Figure 5.2 represents variation of 10% lower and higher than the AVHRR inputs of biomass into the model. Changes in daily methane flux estimates due to differences in biomass are manifested in this model primarily at the peak of the reproductive phase of the rice crop, around day 55 after flooding. This clearly illustrates the fact that the van Bodegom model focuses more on the microbial processes at work in methane emission due to rice crops, and that the growth of the rice plant plays a secondary role. However, the role seems adequate in the sense that the contribution of the rice plant to carbon sources, which comes into play in the later parts of the growing season, result in good estimates of daily

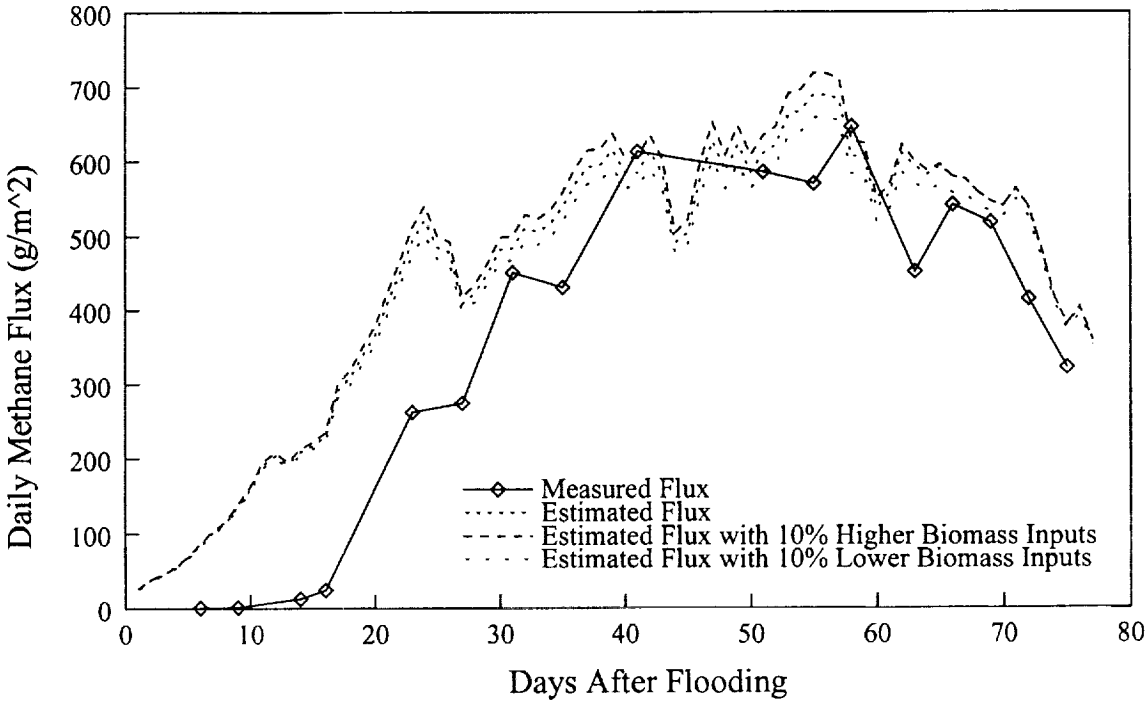



Figure 5.2: Sensitivity of van Bodegom Model to Biomass Inputs

methane flux.

Referring to Figure 5.3, one can note that the changes in daily methane flux for various levels of soil mineralizable carbon content are pronounced throughout almost the entire growing season, with the exception of the start of the season. The reason that rate of methane flux is not affected at the beginning of the season is that the methanogenic bacteria are outcompeted for the substrate production resulting from the mineralization of the carbon by the alternative electron acceptors. Once those acceptors are depleted, the methanogens can utilize the substrates.

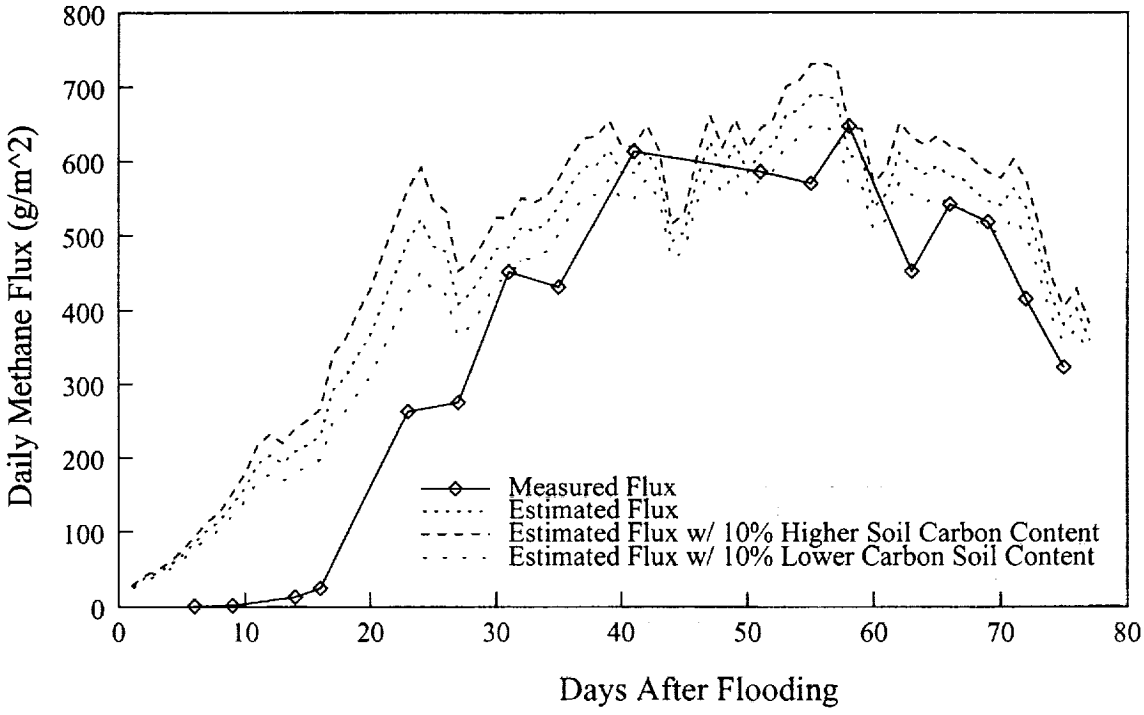



Figure 5.3: Sensitivity of van Bodegom Model to Soil Carbon Content

The effects of varying soil iron content can be ascertained by looking at the plot shown in Figure 5.4. Here is presented not only a +- 10% difference in Fe(III) levels, but also the Fe for a soil very different from the soil found in this region of Texas. The effect on methane flux, as expected, is more pronounced at the start of the growing season. This again relates to the competition between the

methanogens and the alternative electron receptors associated with iron.

A summary of the results from the coupling of AVHRR-derived data with the van Bodegom model will be presented in Chapter 7, as will suggestions for further research.

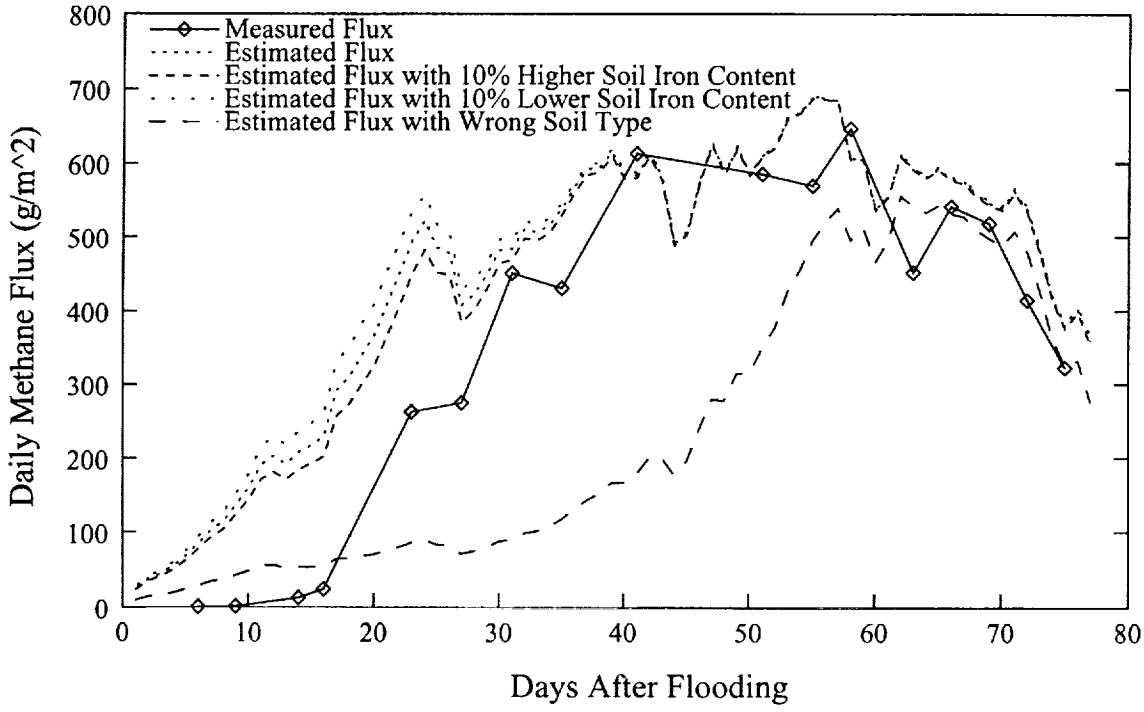



Figure 5.4: Sensitivity of van Bodegom Model to Soil Iron Content

# Chapter 6

# Mapping Rice with AVHRR

## 6.1 Background

In an effort to take advantage of the advances presented in previous chapters, specifically, using satellite derived biomass to drive methane emission estimation models, one can envision some type of GIS database with georeferenced information on soil types and rice crop locations, which would allow for automatic inputs. Information is already available to populate a GIS soils layer in such a database. However, exact locations of rice crops from year to year are pieces of information that are not available from any sources at this time. For example, in the United States, crop information is generally disseminated in terms of acreage, such as number of acres planted/harvest for a given crop. The exact locations of the fields which are planted are of little consequence to the tabulation of such crop statistics by state or federal agencies. Some organizations that do care about exact locations are forced to make such mappings on their own. For example in Texas, personnel in the offices of the Lower Colorado River Authority (LCRA) who deal with supplying water to rice paddies must hand-draw reference maps each year to record which fields are planted and which are left fallow because the information is not available elsewhere.

With these thoughts in mind, a preliminary investigation was conducted to

assess a technique with potential for automatically detecting likely rice growing areas directly from the same satellite imagery used to obtain the NDVI time series-based estimates of biomass for the methane flux prediction models. This technique, called matched filtering, is described in this chapter and the results of its application in this research project are presented and discussed.

## 6.2 Methods and Materials

### 6.2.1 Matched Filtering

Matched filters were originally developed to detect known signals in the presence of noise [9] and have traditionally been used in signal processing applications, such as voice detection/recognition. A logical extension of this was an attempt to detect a known signal within a group of unknown signals. It is in this spirit that the filter has been applied to satellite imagery.

One application of a matched filter to imagery is to utilize known spatial patterns to detect specific targets, if present in the scene, against various backgrounds. In this application, as well as in the signal processing application, the matched filter works by taking advantage of properties of spatial information transformed into the frequency domain. For example, if one applies a Fourier transform to both a known signal and an unknown signal, the convolution of the two will result in a delta function if the two signals are perfectly matched. Peaks of decreasing amplitude and increasing width will result from increasingly dissimilar signals. In this way, one can obtain an indication of how closely the two signals match, or do not match, by the result of the convolution. Figure 6.1 illustrates, conceptually, how the matched filter searches for a reference signal, shown in (a), in a noisy signal, shown in (b), and the output of the matched filter operator, in (c).

(a) Reference (b) Reference+Noise (c) Output

Figure 6.1: Schematic of Matched Filtering Inputs and Output in Frequency Domain[9]

In this study, instead of using a known spectral signature to detect likely rice growing areas, a known temporal signature, in the form of an NDVI time series, is used. Because cropping conditions change from year to year, the reference NDVI time series will also change from year to year. Thus, it would not be possible to build a reliable library which would be applicable to each growing season of interest. Instead, a reference time series, or spectrum, would need to be defined for each year under study.

Unlike the spatial matched filter, the spectral matched filter does not transform the image data into the frequency domain. Instead, the filter performs and considers the results of linear transformations which are commonly used in the analysis of multispectral data sets. Rather than coding a matched filtering algorithm, an existing implementation of this filter was used from the Envi software package available from RSI Inc. The following summary of the mathematical basis used in the Envi software now follows.

According to Harsanyi and Chang [12], it is possible to isolate a signature of interest from undesired signatures through the development of a specially designed linear operator. This operator could then be used on any given mixed pixel, $\mathbf{r}$ of a multispectral image,

$$\mathbf{r}(x,y) = \boldsymbol{M}\boldsymbol{\alpha}(x,y) + \mathbf{n}(x,y) \tag{6.1}$$

where M(x,y) is a matrix with linearly independent columns of $p$ different spectral signatures and rows representing each of the bands under consideration. The $\boldsymbol{\alpha}$(x,y) is a $p$x1 vector with the weights of each spectral signature in the pixel, and $\mathbf{n}$ is a vector representing random noise for each band. If the last column of $\mathbf{M}$ is assumed to be a spectrum of interest, and called $\mathbf{d}$, then (6.1) can be rewritten as

$$\mathbf{r} = \mathbf{d}\alpha_{\mathrm{p}} + \mathbf{U}\boldsymbol{\gamma} + \mathbf{n} \tag{6.2}$$

where $\alpha$ is the fraction of the spectrum of interest, $\mathbf{U}$ is a matrix with linearly independent columns of $p$-1 spectral signatures, and $\boldsymbol{\gamma}$ is a vector which contains the first $p$-1 values of $\boldsymbol{\alpha}$.

In order to develop the matched filter such that it eliminates the effects of the unwanted spectra represented by $\mathbf{U}$, the mixed pixel $\mathbf{r}$ must be projected onto a subspace orthogonal to $\mathbf{U}$. In this way, the resulting vector will contain only energy associated with the spectrum of interest, as well as random noise.

Harsanyi and Chang continue by defining this operator to eliminate the unwanted spectra to be

$$\mathbf{P} = (\mathbf{I} - \mathbf{U}\mathbf{U}^{\#}) \tag{6.3}$$

where $\mathbf{U}^{\#} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T$ is the pseudoinverse of $\mathbf{U}$. As opposed to minimizing the least squares error, this operator instead minimizes the energy associated with the spectra which are not of interest. Using $\mathbf{P}$ to operate on (6.2) yields

$$\mathbf{P}\mathbf{r} = \mathbf{P}\mathbf{d}\alpha_{\mathrm{p}} + \mathbf{P}\mathbf{n} \tag{6.4}$$

The next step in the construction of the matched filter is to construct $\mathbf{x}^T$, an operator which maximizes the signal-to-noise ratio (SNR).

$$\mathbf{x}^T\mathbf{P}\mathbf{r} = \mathbf{x}^T\mathbf{P}\mathbf{d}\alpha_{\mathrm{p}} + \mathbf{x}^T\mathbf{P}\mathbf{n} \tag{6.5}$$

In order to find a value which maximizes $\mathbf{x}^T$ one must maximize the signal-to-noise-energy ratio, $\lambda$

$$\lambda = \frac{\mathbf{x}^T\mathbf{P}\mathbf{d}\alpha_p^2\mathbf{d}^T\mathbf{P}^T\mathbf{x}}{\mathbf{x}^T\mathbf{P}\mathbf{E}\{\mathbf{n}\mathbf{n}^T\}\mathbf{P}^T\mathbf{x}} = \frac{\alpha_p^2}{\sigma^2}\frac{\mathbf{x}^T\mathbf{P}\mathbf{d}\mathbf{d}^T\mathbf{P}^T\mathbf{x}}{\mathbf{x}^T\mathbf{P}\mathbf{P}^T\mathbf{x}} \tag{6.6}$$

where $\mathbf{E}\{.\}$ denotes the expected value. Maximization of the quotient $\lambda$ can be generalized to an eigenvector problem of the form

$$\mathbf{P}\mathbf{d}\mathbf{d}^T\mathbf{P}^T\mathbf{x} = \tilde{\lambda}\mathbf{P}\mathbf{P}^T\mathbf{x} \tag{6.7}$$

where $\tilde{\lambda} = \lambda(\alpha_p^2/\sigma^2)$. The result of the maximization is

$$\mathbf{x}^T = \kappa\mathbf{d}^T \tag{6.8}$$

where $\kappa$ is an arbitrary scalar.

Finally, according to Harsanyi and Chang, substitution of (6.8) into (6.5) yields the overall matched filter operator

$$\mathbf{q}^T = \mathbf{d}^T\mathbf{P}. \tag{6.9}$$

When this matched filter operator is applied to each multiband pixel of an image being searched, a single band image results with the value of each pixel being a measure of the presence of the signature of interest. [12]

Included in Figure 6.2 is a schematic representing the matched filter results for three different pixels. In (a) the NDVI time series of the unknown pixel from an NDVI time series cube, exactly matches the reference in (b) which would result in a relative match value of 1 being assigned to the pixel location of the unknown time series in the matched filter output "image". In (c) is shown an NDVI time series for another unknown pixel in the same image cube. Comparing it again to the reference time series in (d), the matched filter would assign a very low relative match value. In (e) is shown an NDVI time series of a third unknown pixel. When

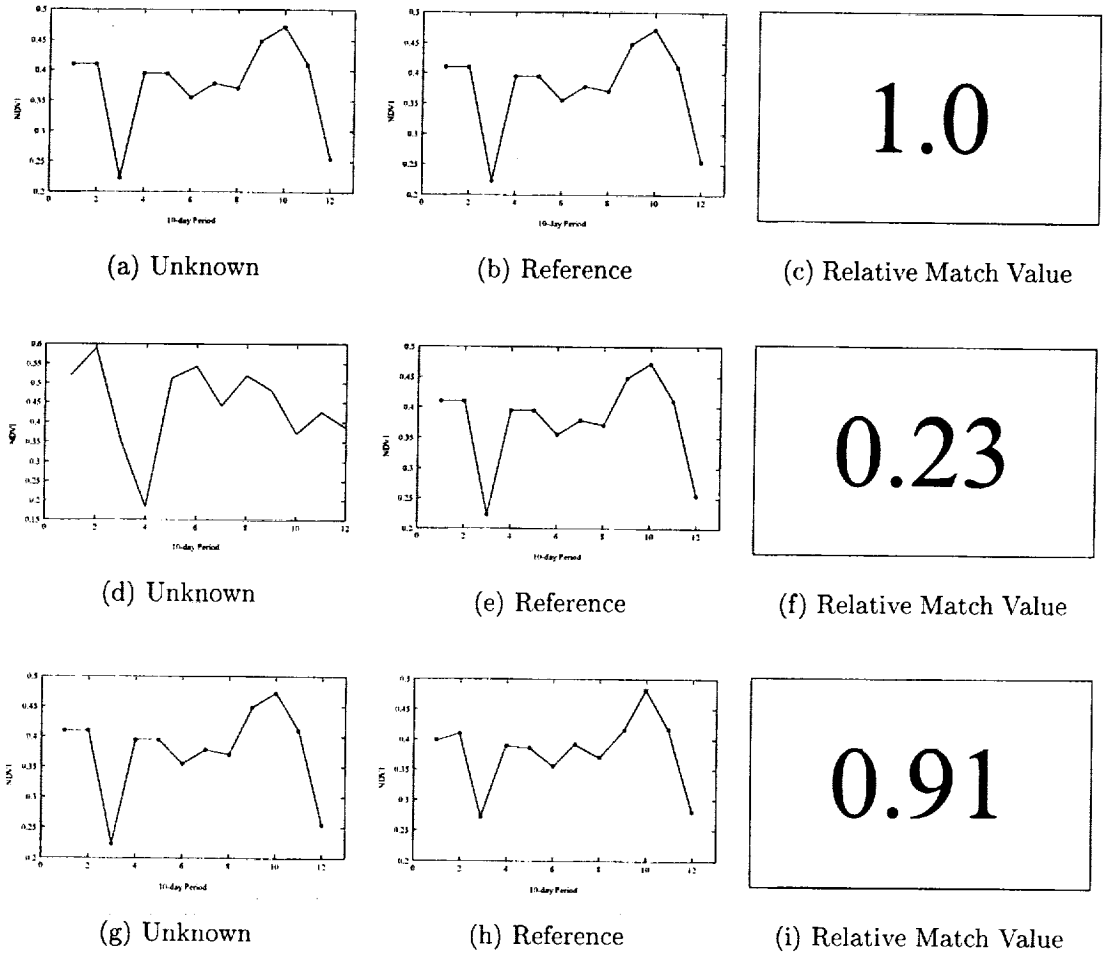
compared to the reference time series, again shown in (f), a fairly high relative match value would be assigned to that pixel location in the matched filter output image since it so similar to the reference.

### 6.2.2 Ground Truth Data

As mentioned in the previous section, one must know for certain at least one area where rice is being grown in order to obtain the reference spectrum. The areas of known rice cropping from which the data used in the previous chapter were obtained, were again used here as the reference spectra for each year under consideration. Again, these locations were at the Texas A&M research station near Beaumont, TX for 1992-1995, and at a farmer's field near El Campo, TX for 1998-9.

As discussed in the previous sections, the output of a matched filter is a value for each pixel which provides a relative degree match between the spectrum of the pixel under consideration and the reference spectrum. Because the values of the pixels in the image resulting from the matched filter can range almost infinitely from 0 to 1 (1 being a perfect match) the challenge is to select an appropriate cutoff for this value, such that false positives are minimized. Because conditions each year are different, as well as the fact that the matched filter result is only relative, we would not expect that the cutoff value would be the same from year to year. In order to guide this cutoff selection, agricultural statistics were utilized to constrain the number of pixels flagged as adequately matching the reference spectrum. These statistics were acquired from the Texas Agricultural Statistics Service (TASS) for the six years under investigation in this study. Figures for the number of acres of rice which were planted for each of the six years are included in Table 6.1, as are the approximate number of AVHRR pixels which represent the area planted with rice.

Once the matched filter was run for each year of interest, a histogram of the resulting degree of match values was produced for each year. The cutoff value for each was selected by noting what degree of match value resulted in the number




(a) Unknown (b) Reference (c) Relative Match Value

(d) Unknown (e) Reference (f) Relative Match Value

(g) Unknown (h) Reference (i) Relative Match Value

Figure 6.2: Schematic of Matched Filtering Inputs and Outputs Using NDVI Time Series

Table 6.1: Acres of Rice Planted and Number of AVHRR Pixels Represented

| Year | Number of Acres Planted in Rice | Number of 1.1km Pixels Representing Acres of Rice |
|---|---|---|
| 1992 | 353,000 | 1180 |
| 1993 | 298,000 | 1000 |
| 1994 | 354,000 | 1190 |
| 1995 | 318,000 | 1065 |
| 1998 | 268,000 | 900 |
| 1999 | 245,000 | 820 |

of pixels expected for rice cropping acreage for that year. Although these cutoff values by themselves are not important, the fact that they ranged from 0.556 - 0.818, reinforces the point that indeed each application of a matched filter yields relative and not absolute results. As an example, the plot shown in Figure 6.3 is the histogram for all of the relative match values for land areas in 1993. It is from histograms like this that the relative match thresholds were selected.

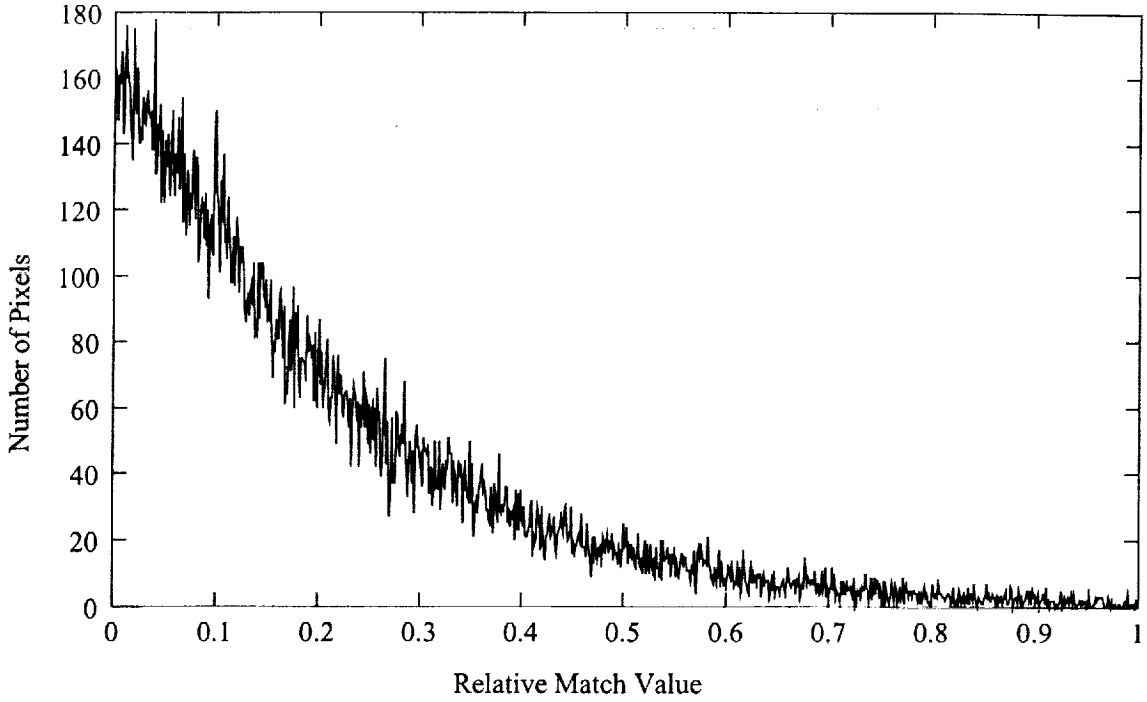
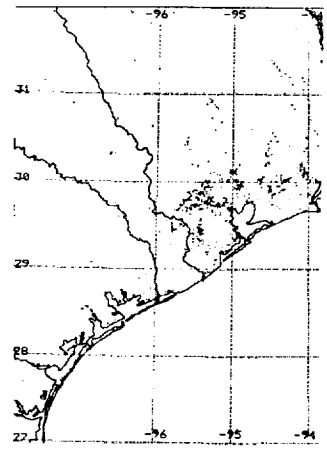



Figure 6.3: Histogram of Relative Match Values for Land Pixels in 1993

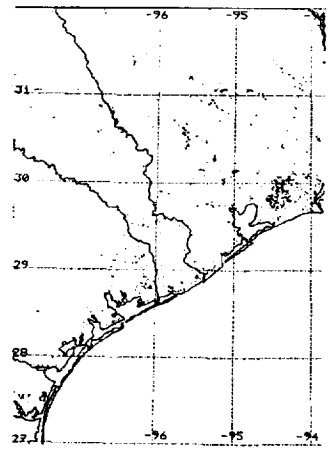
## 6.3 Results and Discussion

The application of the Envi-based matched filter, using the crop statistics thresholding criterion described above, yields mixed results. The accuracy of the method can not be quantitatively measured because, at this time, not enough ground truth is available to do so. However, qualitative assessments can be made and will be described in this section. Suggestions for further assessments with the availability of ground truth, as well as different approaches will be discussed in Chapter 5.

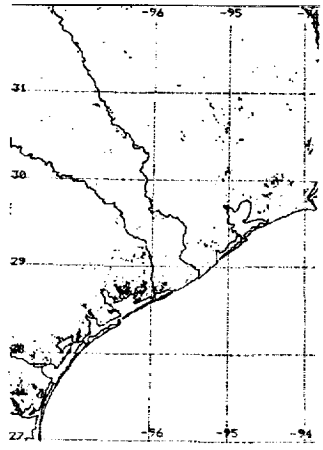
We begin with making some observations about the maps presented in Figure 6.4 on page 61. Each of the maps in the subfigures represents the rice growing region of Texas, and the pixels which met the relative match value threshholds are shown in black. The first observation one can make is that the majority of the pixels which are the likeliest match to the rice NDVI time series are clustered



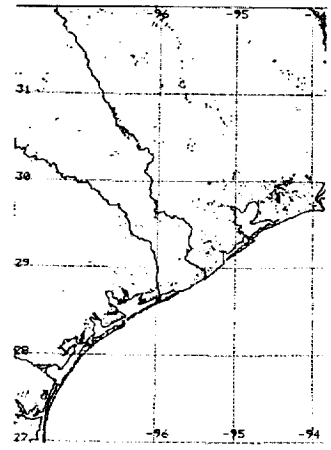
(a) 1992



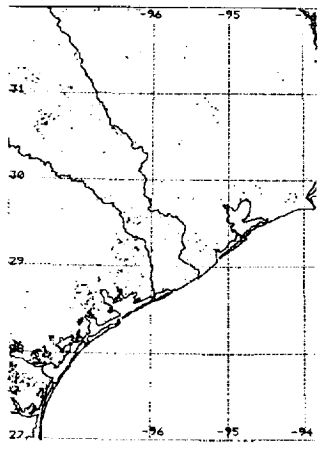
(b) 1993



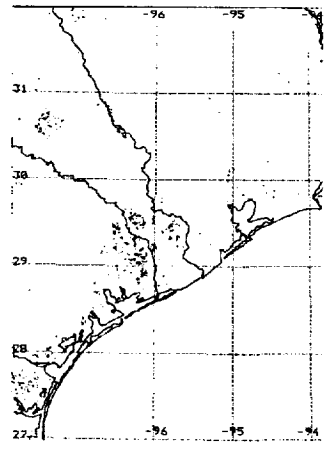
(c) 1994



(d) 1995



(e) 1998



(f) 1999

Figure 6.4: Matched Filter Results of Likely Rice Cropping Areas for 1992-5 and 1998-9

in the coastal areas of the region, which is indeed where the rice is grown. Few, if any, pixels in the upper left region of the maps were selected as having a high degree of match. Since this part of Texas is generally uncultivated grassland, this observation indicates a good amount of promise for the matched filter in this application.

Closer inspection, however, reveals that the locations of the concentration of the matched pixels shift rather dramatically from year to year. If familiar with the reference map shown in Fig. 3.1, one can note which counties have the highest concentrations. The following table summarizes these observations for each year and also provides those counties' rank out of the eighteen rice producing counties in Texas (in terms of acreage planted). A rank of one would denote the highest producing county.

Table 6.2: Counties with Concentrations of Matched Rice Pixels

| Year | Counties with High Concentration of Matched Pixels | Rank of the County as Rice Producer in Texas (of 18) for that year |
|---|---|---|
| 1992 | Harris | 10 |
| | Wharton | 1 |
| | Ft. Bend | 9 |
| 1993 | Jefferson | 3 |
| | Chambers | 6 |
| 1994 | Matagorda | 3 |
| | Jackson | 6 |
| | Non-rice producing | – |
| 1995 | Jefferson | 3 |
| | Chambers | 6 |
| 1998 | Wharton | 1 |
| | Colorado | 2 |
| | Non-rice producing | – |
| 1999 | Wharton | 1 |
| | Colorado | 2 |
| | Non-rice producing | – |

From Table 6.2, one can see that three of the six years had significant numbers of pixels which the matched filter operator identified as having a high likelihood of being rice, but were outside of the rice cropping areas of Texas. Additionally, only a handful of the top rice producing counties consistently had significant numbers of rice cropping pixels attributed to them. So although at first glance, it would appear through visual inspection that the matched filtering operator does a reasonable job in detecting likely rice cropping areas, only a rather cursory qualitative look strongly indicates that this technique, as it is currently implemented, does not yield accurate enough results for the intended purpose.

In trying to assess whether or not the matched filter operator has unrealized potential in utilizing temporal signatures, rather than the more well-proven

spectral and spatial signatures, the weaknesses of its current implementation must be explored in order to see if there are ways to potentially overcome them. The weaknesses can fall into one of three categories: the source imagery of the temporal signature, the selection criterion for the relative match value cutoff, and the use of the temporal signature itself.

In assessing the source imagery of the temporal signature, namely the AVHRR sensor, one must consider the attributes of the sensor itself, as well as the feasible processing that can be done for large ground coverage applications such as this. In terms of the sensor, the rather gross spatial resolution is clearly an issue. Although the matched filter operator is designed to be applied to cases of mixed pixels, the relative contribution of a spectrum of interest has a higher potential to be occulted in a pixel covering a larger ground area. Because of the greater potential for mixed spectra, even pixels having mid-range relative match values due to the presence of the spectrum of interest would not be likely to have a high enough match value to pass the cutoff threshold, and would thus be eliminated from any final count. Obviously a pixel which has a smaller footprint would have greater potential for allowing the spectrum of interest to differentiate itself from other spectra.

The relatively recent launch of the AVHRR follow-on mission, the Moderate Resolution Imaging Spectroradiometer (MODIS) will go a long way in improving on the spatial resolution of NDVI products (from 1.1 km for AVHRR to 250 m for MODIS) while maintaining the superior temporal resolution. This improvement will help to accomplish goals such as the development of image segmentation operators which can be applied to tasks which require increased levels differentiation among vegetation types.

The issue of which pre-processing techniques are used in this study for the

AVHRR imagery may also have an impact on the successful application of the matched filter operator. Because this operator maximizes differences between the reference spectrum and the other spectra in the image, pixels which actually have a similar temporal spectrum to the reference, but which have even one ten-day period in which the pixel was cloud-covered, will result in a low degree of match for that pixel. Therefore, a method of reconstructing missing data values in an NDVI time series which are due to clouds would need to be found in order to increase the chances that the pixel would be correctly identified. Proposed in the recent literature is a new technique for doing so which uses the Fourier transform of the NDVI time series to detect the clouds, and then utilizes a statistical method for reconstruction of any missing portions of the time series [30]. Successful application of such a method would allow the matched filter to detect more true rice growing areas since the threshold would no longer eliminate those pixels in favor of some false positives.

The issue of the selection criterion for the cutoff value for the relative match value could also contribute to the less than desirable results which were obtained in this application of the matched filter operator. The idea of guiding classifications with agricultural statistics has been explored with some success in the past (see [47]) so its use here is not being questioned. However, the way in which the statistics were utilized in this application may be able to be improved. In this study, since it is known that rice cropping is done only in the Gulf coast area of Texas, imagery from the entire state was not included. However, significant areas where it is known that rice is not grown were still included in the input to the matched filter operator. It is conceivable that a GIS system could be utilized to build a mask to eliminate those counties which are not involved in rice cropping, such that the matched filter operator is applied only to those counties which are.

Then, the agricultural statistics could still be utilized in the same manner to select the relative match value threshold, presumably with better results.

Finally, the issue of the efficacy of using a temporal reference spectrum for rice from one region and expecting it to adequately characterize rice in another region should be addressed. As discussed previously, there were two ground truth sites in two different parts of the rice growing region of Texas. As shown in the Soils Table presented in Chapter 3, these sites are situated on two distinctly different soil types. One can see in Figure 6.4 on page 61 that the 1992-1995 matched filter outputs did not effectively highlight the rice cropping areas around El Campo, while the 1998-1999 matched filter outputs did not effectively highlight the rice cropping areas around Beaumont. A question that should be asked is if it is possible that these soil differences, and the differences in cropping practices which may result from the difference in plant requirements, can cause enough differences in the rice plants that the NDVI time series for the two are too different to be used to detect the other. In an effort to better understand this, a closer look was given to the NDVI time series, shown here in Figure 6.5, for each of these growing seasons.

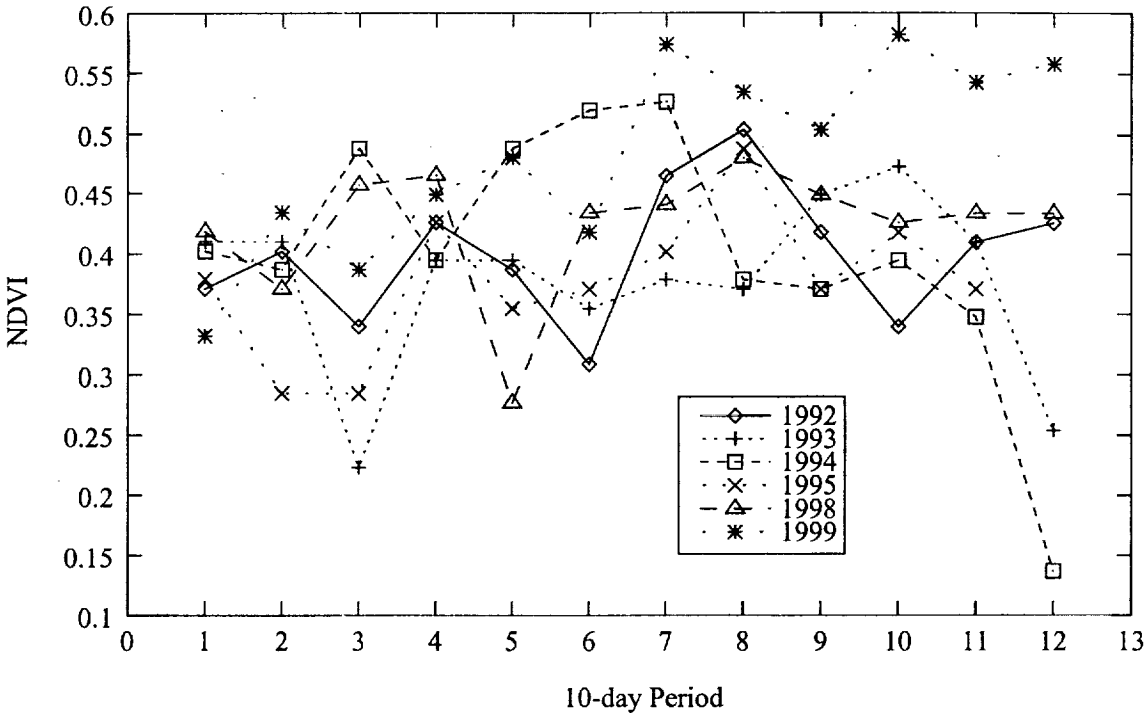



Figure 6.5: NDVI Time Series for 1992-1995 and 1998-1999

In trying to use the information in the plot to support the idea that the NDVI time series are just too different for one region to use in another region, one would expect noticeable differences between the 1992-1995 Beaumont spectra and the 1998-1999 El Campo spectra. Instead of detecting noticeable differences between the two groups, one can note the each of the NDVI time series presented has unique features. There seems to be no Beaumont grouping and no El Campo grouping. Upon further consideration, because the matched filter operator responds to relative factors within an image cube, one can not expect that differences in results from case to case can be tied to differences among their respective reference spectra.

# Chapter 7

# Summary and Conclusions

## 7.1 Coupling AVHRR Data with Methane Emission Models

In this research, the promise of coupling satellite remote sensing image data with biogeochemical models of methane from rice was demonstrated. The crux of the demonstration was relating a satellite image derived value, NDVI, with a plant parameter, biomass, that could be used to drive the models. The empirical relationship between cumulative NDVI and total aboveground biomass provides a satisfying approximation for biomass that also has some rationale from a physical standpoint.

The limitations of the relationship that was derived are, however, not negligible. First among them is the fact that the study area was limited to Texas, and the rice varieties and growing season are not likely to be directly applicable to areas outside of the Gulf of Mexico region of Texas, and perhaps Louisiana. If future study in other geographic areas is undertaken, one priority should certainly be an attempt to identify sources of plant data in other regions, such as California, and of course, in Southeast Asia. If such data were acquired, and it is found that indeed, the cumulative NDVI-to-biomass relationship derived in this study does not hold, then perhaps a limited database of relationships could be developed such that future analysis could select the relationship appropriate to

the region under study.

One difficulty with utilizing an accumulating variable, in this case, cumulative NDVI, is that changes in lengths of growing seasons over the compositing rate (in this case, 10 days) render the relationship unusable. This problem was already encountered in the course of the research performed for this study, as will be described here. An excellent source of methane flux data was acquired under the direction of Dr. M.A.K. Khalil over seven years in the Sichuan Province of China [19]. With the view to utilize this data set, AVHRR imagery was acquired for three of the seven years that Khalil collected methane flux data. It was assumed that the cumulative NDVI-to-biomass relationship derived and described in Chapter 3 could be utilized to estimate the biomass of the Chinese rice using the NDVI time series of the China imagery. However, it was discovered that the length of the growing season for rice was over 20 days longer in this study area of China than in the Texas study areas. Thus, the cumulative NDVI-to-biomass relationship would use two extra NDVI MVC composites in accumulation, producing biomass values that would likely be much higher than in reality. Assuming that the biomass values for the Chinese varieties are similar to the Texas varieties, the plants simply take longer to grow. Thus, a cumulative NDVI to biomass relationship for the China area would have a considerably lower slope than the relationship derived for Texas.

Another difficulty that would certainly be posed by forming a cumulative-based relationship will occur in regions where more than one rice crop is grown per year, such as in the southern provinces of China where up to three rice crops are grown per year. Because of this, the accumulation of NDVI over the whole year would cease to be linear.

Assuming that the methodology of obtaining biomass using cumulative

NDVI from AVHRR can be successfully applied to areas other than Texas, the use of the Huang model, and even the van Bodegom model with improved access to accurate inputs, show great promise in being able to estimate methane emission, both on a daily flux basis and on a total seasonal flux basis. Eliminating the dependency of obtaining a yield value for estimating biomass, used in both the Huang model and the original van Bodegom model, is an important result of this research. Crop yield values, although generally easy to obtain, can vary tremendously. National- and state-level yield values are averaged to such a great extent, that it would be unwise to rely on them at the county- and certainly not at the pixel level. Obtaining county-level yields are much more difficult, and often rely on voluntary information provided by individual farmers. Estimating biomass from satellite imagery instead of yield seems like a much better method, and one which, as shown in this research, has the potential to result in as good an estimate of methane emission from rice as areas where yield information is well known.

Although this study was not intended to directly compare the results of the two models, it is appropriate to discuss the implication that the Huang model estimated methane emission better than did the van Bodegom model. This conclusion is the result of comparing average daily methane fluxes and total seasonal fluxes obtained from the two models with those fluxes which were measured. The flux estimates obtained from the Huang model seemed to be more closely matched to the measured values than did the van Bodegom estimates. However, caution should be made when making these types of comparisons. Because the methane flux measurements were made on only 17 of the 77 days that the paddy was flooded, there may be an over- or underestimation of the daily average and total seasonal fluxes resulting from small time-scale flux changes being missed through undersampling. Additionally, the greater difference between the van Bodegom es-

timates and the measured values may be due to imprecisions in the non-biomass inputs, especially the mineralizable carbon and iron contents of the soil. An additional item to note is that the Huang model, since it is more empirical has likely been better tuned to this study area. As a result, the model may not perform as well in other rice cropping areas. Finally, it should be pointed out that there was only one year of methane flux data available. Additional years of data, with different cropping conditions, would better provide understanding of the strengths and limitations of each of the models in relation to using AVHRR derived biomass inputs.

This brings up the final summary point, which is that although one of the major ground truth inputs, namely crop yield data, has been able to be replaced with a satellite derived input, each model still requires other inputs. Some of the inputs, such as average air temperatures, are fairly easy to obtain. Others, such as soil components, are not. Still others, such as variety index and specific root length, may only be estimated because the information does not exist at all. Work must continue to create better and more reliable databases of the necessary information which already exists in other places, and to work with scientists who are in a better position to populate the databases for information which does not yet exist.

## 7.2 Mapping Rice with AVHRR

Because of the promise demonstrated in this research to drive methane flux estimation models with biomass of rice plants derived from satellite data, a means to identify the locations of rice paddies from year-to-year would be invaluable. The matched filter method explored in this research is not up to that task in its present form and implementation. But the idea should not be given up entirely. Several

suggestions for modifications to the manner in which the filter is implemented were discussed extensively in the previous chapter. These include constraining the image area searched using crop statistics.

Another feature that should be added to the information that the mapping routine provides is growth season-related estimates, such as when the rice crop started, when it was harvested, etc. These types of dates are required for the implementation of the satellite derived biomass in the methane emission estimation models. Although the rice planted in each region likely maintains the same planting-to-harvest duration, the time at which planting occurs varies from year to year, based on weather conditions. In order to accumulate the proper NDVI time series, the start of the cropping season needs to be known. This type of information should be determined at the same time that the rice cropping areas are being detected.

Overall, the potential for increased understanding of the contribution of rice paddies to the atmospheric methane budget through the use of AVHRR satellite data has been demonstrated. Future research and study will help to realize the potential, and would be time and money well spent.

## Bibliography


[1] D. Bachelet. Rice paddy inventory in a few provinces of China using AVHRR data. Geocarto International, 10(1):23–38, 1995.

[2] D. Bachelet and H.U. Neue. Methane emissions from wetland rice areas of Asia. Chemosphere, 26(4):219–237, 1993.

[3] Th. A. de Boer. Botanical characteristics of vegetation and their influence on remote sensing. In Henk J. Buiten and Jan G.P.W. Clevers, editors, Land Observation by Remote Sensing: Theory and Applications, pages 89–106. Gordon and Breach Science Publishers, Amsterdam, 1993.

[4] Guy Brasseur and David Schimel. Atmospheric chemistry and the earth system. In Guy Brasseur, John J. Orlando, and Geoffrey S. Tyndall, editors, Atmospheric Chemistry and Global Change, Topics In Environmental Chemistry. Oxford University Press, New York, 1999.

[5] Mingkui Cao and J.B. Dent. Modeling methane emissions from rice paddies. Global Biogeochemical Cycles, 9(2):183–195, 1995.

[6] David J. Carty. Characteristics of pimple mounds associated with the Morey Soil in Southeast Texas. M.S., Texas A&M University, 1980.

[7] D. Casanova, G.F. Epema, and J. Goudriaan. Monitoring rice reflectance at field level for estimating biomass and LAI. Field Crops Research, 55:83–92, 1998.

[8] Kenneth R. Castleman. Digital Image Processing. Prentice-Hall Signal Processing Series. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1979.

[9] William R. Cotton and Roger A. Pielke. Human Impacts on Weather and Climate. Cambridge University Press, Cambridge, 1995.

[10] R.S. DeFries and J.R.G. Townshend. NDVI-derived land cover classifications at a global scale. International Journal of Remote Sensing, 15(17):3567–3586, 1994.

[11] Joseph C. Harsanyi and Chein-I Chang. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. IEEE Transactions on Geoscience and Remote Sensing, 32(4):779–785, 1994.

[12] David A. Hastings and William J. Emery. The advanced very high resolution radiometer: A brief refence guide. Photogrammetic Engineering & Remote Sensing, 58(8):1183–88, 1992.

[13] A. Holzapfel-Pschorn, R. Conrad, and W. Seiler. Effects of vegetation on the emission of methane from submerged paddy soil. Plant and Soil, 92:223–33, 1986.

[14] Yao Huang. Methane Emission from Irrigated Rice Cultivation: Quantities, Models, and Practice. Ph.D., Rice University, 1997.

[15] Yao Huang, Ronald L. Sass, and Frank M. Fisher. Methane emission from texas rice paddy soils. 1. quantitative multi-year dependence of CH4 emission on soil, cultivar and grain yield. Global Change Biology, 3:479–489, 1997.

[16] Yao Huang, Ronald L. Sass, and Frank M. Fisher. A semi-empirical model of methane emission from flooded rice paddy soils. Global Change Biology, 4:247–268, 1998.

[17] M.A.K. Khalil. An introduction to atmospheric methane. In M.A.K. Khalil, editor, Atmospheric Methane: Sources, Sinks, and Role in Global Change, NATO ASI Series, pages 1–6. Springer-Verlag, Berlin, 1993.

[18] M.A.K. Khalil, R.A. Rasmussen, M.J. Shearer, R.W. Dalluge, L.X. Ren, and C.-L. Duan. Measurements of methane emissions from rice fields in China. Journal of Geophysical Research, 103(19):25181–25210, 1998.

[19] Brigitte Leblon, Martine Guerif, and Frederic Baret. The use of remotely sensed data in estimation of PAR use efficiency and biomass production of flooded rice. Remote Sensing of the Environment, 38:147–158, 1991.

[20] Thomas M. Lillesand and Ralph W. Kiefer. Remote Sensing and Image Interpretation. John Wiley & Sons, New York, 2nd edition, 1987.

[21] Sietse O. Los, Christopher O. Justice, and Compton J. Tucker. A global 10by10 NDVI data set for climate studies derived from the GIMMS continental NDVI data. International Journal of Remote Sensing, 15(17):3493–3518, 1994.

[22] William B. Meyer, W. Neil Adger, Katrina Brown, Dean Graetz, Peter Gleick, John F. Richards, and Antonio Maghalaes. Land and water use. In Steve Rayner and Elizabeth L. Malone, editors, Human Choice and Climate Change: Resources and Technology, volume 2, pages 79–134. Battelle Press, Columbus, Ohio, 1998.

[23] Victor I. Myers. Crops and soils. In Leonard W. Bowden, editor, Manual of Remote Sensing, volume 2, pages 1715–1727. American Society of Photogrammetry, Falls Church, Virginia, 1975.

[24] Heinz-Ulrich Neue. Methane emission from rice fields. BioScience, 43(7):466–474, 1993.

[25] H.U. Neue and P.A. Roger. Potential of methane emission in major rice ecologies. In Richard G. Zepp, editor, Climate-Biosphere Interactions : Biogenic Emissions and Environmental Effects of Climate Change, pages 65–94. Wiley, New York, 1994.

[26] K.W. Oleson, S. Sarlin, J. Garrison, S. Smith, J.L. Privette, and W.J. Emery. Unmixing multiple land-cover type reflectances from coarse spatial resolution satellite data. Remote Sensing of the Environment, 54:98–112, 1995.

[27] John C. Price and Water C. Bausch. Leaf area index estimation from visible and near-infrared reflectance data. Remote Sens. Environ., 52(1):55–65, 1995.

[28] H. Rahman and G. Dedieu. SMAC: A simplified method for the atmospheric correction of satellite measurements in the solar spectrum. Int. J. Remote Sensing, 15(1):123–143, 1994.

[29] G.J. Roerink and M. Menenti. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing, 21(9):1911–1917, 2000.

[30] J.A Romberger, Z. Hejnowicz, and J.F. Hill. Plant Structure: Function and Development. Springer-Verlag, Berlin, 1993.

[31] Steven W. Running. Estimating terrestrial primary productivity by combining remote sensing and ecosystem simulation. In R.J. Hobbs and H.A. Mooney, editors, Remote Sensing of Biosphere Functioning, volume 79 of Ecological Studies: Analysis and Synthesis, pages 65–86. Springer-Verlag, New York, 1990.

[32] Steven W. Running, Thomas R. Loveland, Lars L. Lierce, R.R. Nemani, and E.R. Hunt Jr. A remote sensing based vegetation classfication logic for global land cover analysis. Remote Sensing of the Environment, 51:39–48, 1995.

[33] Helmut Schuetz, Peter Schroeder, and Heinz Rennenberg. Role of plants in regulating the methane flux to the atmosphere. In Thomas D. Sharkey, Elizabeth A. Holland, and Harold A. Mooney, editors, Trace Gas Emissions by Plants. Academic Press, Inc., San Diego, 1991.

[45] M. M. Verstraete and S. Flasse. Recent advances in algorithm development to extract information from AVHRR data. In Giles D'Souza, Alan S. Belward, and Jean-Paul Malingreau, editors, Advances in the Use of NOAA AVHRR Data for Land Applications, volume 5 of EuroCourses, pages 211–229. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1 edition, 1996.

[46] Paul A. Walker and Thilak Mallawaarachchi. Disaggregating agricultural statistics using NOAA AVHRR NDVI. Remote Sensing of the Environment, 63:112–125, 1998.

[47] R. Wassmann, H. Paper, and H. Rennenberg. Methane emission from rice paddies and possible mitigation strategies. Chemosphere, 26(4):201–217, 1993.

[48] Guowei Wu. Biological Systems Approach to Analyzing Population Dynamics of Rice and its Response to Major Physical and Biotic Variables. Ph.D., Texas A&M University, 1996.

[49] Donald J. Wuebbles and John S. Tamaresis. The role of methane in the global environment. In M.A.K. Khalil, editor, Atmospheric Methane: Sources, Sinks, and Role in Global Change. Springer Verlag, Berlin, 1991.